

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 897.273 Seconds  
(without alignments)  
1504.579 Million cell updates/sec

Title: US-10-082-772-1  
Perfect score: 33  
Sequence: 1 gcrctagaccacatgggaaggcgaagtca 33

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22.4	67.9	100808	9	AC108081 Homo sapi
C 2	22.4	67.9	141208	2	AF001492 Homo sapi
C 3	22.4	67.9	168426	2	AC116783 Mus muscu
C 4	22.4	67.9	186768	10	AC127560 Mus muscu
C 5	22.4	67.9	195637	2	AF002420 Homo sapi
C 6	22.4	67.9	200557	9	AC009052 Homo sapi
C 7	22.4	67.9	203240	2	AC136859 Rattus no
C 8	22.4	67.9	228999	2	AC119491 Rattus no
C 9	22.4	67.9	239681	2	AC128835 Rattus no
C 10	22	66.7	935	8	AY086223 Arabidops
C 11	22	66.7	37570	8	ATU53501 Arabidops
C 12	22	66.7	109694	8	F23A5 Arabidops
C 13	22	66.7	175057	10	AL671893 Mouse DNA
C 14	22	66.7	218976	2	AC120853 Mus muscu
C 15	22	66.7	240155	2	AC139751 Mus muscu
C 16	21.8	66.1	39443	9	AC010505 Homo sapi
C 17	21.8	66.1	216261	2	AC117053 Rattus no
C 18	21.4	64.8	141832	9	AF000432 Homo sapi
C 19	21.4	64.8	196055	2	AC025513 Homo sapi
C 20	21.4	64.8	340000	9	AF001670 Homo sapi
C 21	21	63.6	640	11	BV060941 S212P6766
C 22	21	63.6	121112	9	AC124319 Homo sapi
C 23	21	63.6	134463	9	AP005205 Homo sapi
C 24	21	63.6	195718	10	AL807252 Mouse DNA
C 25	21	63.6	245267	2	AC106568 Rattus no
C 26	21	63.6	270618	2	AC133365 Rattus no
C 27	20.8	63.0	18272	10	AF263366 Mus muscu
C 28	20.8	63.0	72290	2	AC023253 Homo sapi
C 29	20.8	63.0	109370	9	AC005414 Homo sapi
C 30	20.8	63.0	111039	9	AL133230 Human DNA
C 31	20.8	63.0	112562	2	AC135829 Rattus no
C 32	20.8	63.0	122272	9	AL359845 Human DNA
C 33	20.8	63.0	162618	2	AC134046 Oryza sat
C 34	20.8	63.0	187737	9	AC007297 Homo sapi
C 35	20.8	63.0	216266	2	AC073797 Mus muscu
C 36	20.8	63.0	226754	10	AL645625 Mouse DNA
C 37	20.8	63.0	229561	2	AL732445 Homo sapi
C 38	20.8	63.0	235151	2	AC125640 Rattus no
C 39	20.8	63.0	236685	2	AC084744 Mus muscu
C 40	20.8	63.0	238337	2	AC127777 Rattus no
C 41	20.8	63.0	243539	2	AC116186 Rattus no
C 42	20.8	63.0	250386	2	AC096198 Rattus no
C 43	20.8	63.0	277226	2	AC105691 Rattus no
C 44	20.8	63.0	277282	2	AC137423 Rattus no
C 45	20.8	63.0	317440	2	AC094294 Rattus no

# ALIGNMENTS

RESULT 1  
AC108081/c 100808 bp DNA linear PRI 04-MAR-2002  
LOCUS Homo sapiens chromosome 16 clone CTC-661116, complete sequence.  
DEFINITION AC108081  
ACCESSION AC108081  
VERSION AC108081.2 GI:19073801  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 100808)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 16  
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 100808)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (25-JAN-2002) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 100808)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (04-MAR-2002) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Mar 4, 2002 this sequence version replaced gi:18369927.  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality  
estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
GenBank flat file format but are available as part  
of this entry's ASN.1 file.  
-----  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality  
estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
GenBank flat file format but are available as part  
of this entry's ASN.1 file.  
-----  
FEATURES  
source Location/Qualifiers  
1..100808  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosomes="16"  
/clone="CTC-661116"  
BASE COUNT 26119 a 24082 c 23816 g 26791 t  
ORIGIN  
Query Match 67.9%; Score 22.4; DB 9; Length 100808;  
Best Local Similarity 81.2%; Pred. No. 29;  
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 GCTCTAGACCACCATGGGAAGGAGGGAAGTC 32  
|||||  
Db 93187 GCTCTAGACCACCATGGGAAGGAGGGAAGTC 93156  
|||||  
RESULT 2  
AP001492/c 141208 bp DNA linear HTG 30-MAY-2000  
LOCUS  
DEFINITION Homo sapiens chromosome 18 clone RP11-774P10 map 18q11.2, WORKING  
DRAFT SEQUENCE, 37 unordered pieces.  
ACCESSION AP001492  
VERSION AP001492.2 GI:8117350  
KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 141208)  
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.  
TITLE Homo sapiens 141,208 genomic DNA of 18q11.2  
JOURNAL Published Only in Database (2000)  
REFERENCE 2 (bases 1 to 141208)  
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (17-MAR-2000) Masahira Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

Kitasato Univ., 1-15-1 Kitasato, Sagami-hara, Kanagawa 228-8555,  
Japan (E-mail:hattori@gsc.riken.go.jp,  
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,  
Fax:81-42-778-9924)  
On May 30, 2000 this sequence version replaced gi:7288186.  
----- Genome Center  
Center: RIKEN Genomic Sciences Center (GSC)  
Center code: RIKEN  
Web site: http://hgp.gsc.riken.go.jp/  
Contact: hattori@gsc.riken.go.jp  
----- Project Information  
Center project name: HumDraft18  
Center clone name: RP11-774P10  
----- Summary Statistics  
Sequencing vector: PCR products; 100% of reads  
Chemistry: Dye-terminator ET-amersham; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 117079 bases at least Q40  
Consensus quality: 127340 bases at least Q30  
Consensus quality: 133354 bases at least Q20  
Insert size: 137608; sum-of-contigs  
Quality coverage: 4.13X in Q20 bases; sum-of-contigs  
-----  
NOTE: This is a 'working draft' sequence. It currently consists of  
37 contigs. The true order of the pieces is not known and their  
order in this sequence record is arbitrary. Gaps between the  
contigs are represented as runs N, but the exact sizes of the gaps  
are unknown. This record will be updated with the finished sequence  
as soon as it is available and the accession number will be  
preserved  
1  
13763 contig of 13763 bp in length  
13864 25815 contig of 11952 bp in length  
25916 31490 contig of 5575 bp in length  
31591 39624 contig of 8034 bp in length  
39725 48843 contig of 9119 bp in length  
48944 56986 contig of 8043 bp in length  
57087 65233 contig of 8147 bp in length  
65334 70006 contig of 4673 bp in length  
70107 74760 contig of 4654 bp in length  
74861 80752 contig of 5892 bp in length  
80853 84662 contig of 3810 bp in length  
84753 88609 contig of 3847 bp in length  
88710 91465 contig of 2756 bp in length  
91566 95774 contig of 4209 bp in length  
95875 99281 contig of 3407 bp in length  
99382 102524 contig of 3143 bp in length  
102625 106160 contig of 3536 bp in length  
106281 109992 contig of 3732 bp in length  
110093 112321 contig of 2229 bp in length  
112422 114489 contig of 2068 bp in length  
114590 116501 contig of 1912 bp in length  
116602 118149 contig of 1548 bp in length  
118250 120828 contig of 2579 bp in length  
120929 122434 contig of 1506 bp in length  
122535 123688 contig of 1154 bp in length  
123789 125498 contig of 1710 bp in length  
125599 126961 contig of 1363 bp in length  
127062 129064 contig of 2003 bp in length  
129185 130302 contig of 1138 bp in length  
130403 131614 contig of 1212 bp in length  
131715 133327 contig of 1613 bp in length  
133428 134656 contig of 1229 bp in length  
134757 136430 contig of 1674 bp in length  
136531 137602 contig of 1072 bp in length  
137703 138705 contig of 1003 bp in length  
138806 140091 contig of 1286 bp in length  
140192 141208 contig of 1017 bp in length  
Sequence updated (26-May-2000).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 37 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
13763: contig of 13763 bp in length  
13764: gap of 100 bp  
13863: contig of 11952 bp in length  
13864: gap of 100 bp  
25915: contig of 5575 bp in length  
25916: gap of 100 bp  
31490: contig of 8034 bp in length  
31491: gap of 100 bp  
31591: contig of 100 bp  
31592: gap of 100 bp  
39624: contig of 100 bp  
39625: gap of 100 bp  
39724: contig of 9119 bp in length  
39725: gap of 100 bp  
48843: contig of 8043 bp in length  
48844: gap of 100 bp  
56985: contig of 8147 bp in length  
56986: gap of 100 bp  
57087: contig of 8147 bp in length  
57088: gap of 100 bp  
65234: contig of 100 bp  
65235: gap of 100 bp  
65334: contig of 4673 bp in length  
65335: gap of 100 bp  
70007: contig of 4654 bp in length  
70008: gap of 100 bp  
74760: contig of 4654 bp in length  
74761: gap of 100 bp  
74861: contig of 5892 bp in length  
74862: gap of 100 bp  
80753: contig of 3810 bp in length  
80754: gap of 100 bp  
84662: contig of 3847 bp in length  
84663: gap of 100 bp  
84763: contig of 3847 bp in length  
84764: gap of 100 bp  
88610: contig of 2756 bp in length  
88611: gap of 100 bp  
91465: contig of 4209 bp in length  
91466: gap of 100 bp  
91566: contig of 3407 bp in length  
91567: gap of 100 bp  
95774: contig of 3407 bp in length  
95775: gap of 100 bp  
95875: contig of 3407 bp in length  
95876: gap of 100 bp  
99281: contig of 3143 bp in length  
99282: gap of 100 bp  
102524: contig of 3143 bp in length  
102525: gap of 100 bp  
102624: contig of 3536 bp in length  
102625: gap of 100 bp  
106160: contig of 3732 bp in length  
106161: gap of 100 bp  
106266: contig of 2229 bp in length  
106267: gap of 100 bp  
109993: contig of 2229 bp in length  
109994: gap of 100 bp  
110093: contig of 2068 bp in length  
110094: gap of 100 bp  
112322: contig of 1912 bp in length  
112323: gap of 100 bp  
112422: contig of 1548 bp in length  
112423: gap of 100 bp  
114489: contig of 1548 bp in length  
114490: gap of 100 bp  
114589: contig of 1548 bp in length  
114590: gap of 100 bp  
116501: contig of 1506 bp in length  
116502: gap of 100 bp  
116602: contig of 1506 bp in length  
116603: gap of 100 bp  
118149: contig of 1154 bp in length  
118150: gap of 100 bp  
118249: contig of 1154 bp in length  
118250: gap of 100 bp  
120828: contig of 1710 bp in length  
120829: gap of 100 bp  
120939: contig of 1363 bp in length  
120940: gap of 100 bp  
122434: contig of 2003 bp in length  
122435: gap of 100 bp  
122535: contig of 1138 bp in length  
122536: gap of 100 bp  
123688: contig of 1138 bp in length  
123689: gap of 100 bp  
123788: contig of 1138 bp in length  
123789: gap of 100 bp  
125498: contig of 1138 bp in length  
125499: gap of 100 bp  
125599: contig of 1138 bp in length  
125600: gap of 100 bp  
126962: contig of 1138 bp in length  
126963: gap of 100 bp  
127062: contig of 1138 bp in length  
127063: gap of 100 bp  
129065: contig of 1138 bp in length  
129066: gap of 100 bp  
130302: contig of 1138 bp in length  
130303: gap of 100 bp  
130403: contig of 1212 bp in length  
130404: gap of 100 bp  
131614: contig of 1212 bp in length  
131615: gap of 100 bp  
131715: contig of 1613 bp in length  
131716: gap of 100 bp  
133327: contig of 1613 bp in length  
133328: gap of 100 bp  
133428: contig of 1229 bp in length  
133429: gap of 100 bp  
134656: contig of 1674 bp in length  
134657: gap of 100 bp  
134757: contig of 1674 bp in length  
134758: gap of 100 bp  
136431: contig of 1072 bp in length  
136432: gap of 100 bp  
136531: contig of 1072 bp in length  
136532: gap of 100 bp  
137603: contig of 1003 bp in length  
137604: gap of 100 bp  
137703: contig of 1003 bp in length  
137704: gap of 100 bp  
138705: contig of 1003 bp in length  
138706: gap of 100 bp  
138805: contig of 1003 bp in length  
138806: gap of 100 bp

\* 138806 140091: contig of 1286 bp in length  
\* 140092 140191: gap of 100 bp  
\* 140192 141208: contig of 1017 bp in length.  
FEATURES  
source  
1. 141208  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18q11.2"  
/clones="RP11-774P10"  
1. 13763  
/note="assembly\_fragment clone\_end:T7 vector\_side:right"  
misc\_feature  
13864..25815  
/note="assembly\_fragment"  
25916..31490  
/note="assembly\_fragment"  
31591..39624  
/note="assembly\_fragment clone\_end:SP6 vector\_side:left"  
misc\_feature  
39725..48843  
/note="assembly\_fragment"  
48944..56986  
/note="assembly\_fragment"  
57087..65233  
/note="assembly\_fragment"  
65334..70006  
/note="assembly\_fragment"  
70107..74760  
/note="assembly\_fragment"  
74861..80752  
/note="assembly\_fragment"  
80853..84662  
/note="assembly\_fragment"

Query Match 67.9%; Score 22.4; DB 2; Length 141208;  
Best Local Similarity 81.2%; Pred. No. 29;  
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GCTCTAGACCATGGGAGAGCGGAGTC 32  
|||||||  
Db 67769 GCTCTAGACCATGGGAGAGCGGAGTC 67738

## RESULT 3

AC116783  
LOCUS  
DEFINITION Mus musculus clone RP24-145D9, WORKING DRAFT SEQUENCE, 7 ordered  
pieces.  
AC116783  
VERSION AC116783.4 GI:28604041  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.

## REFERENCE

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## 1 (bases 1 to 168426)

## AUTHORS

Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Reback, M., Riley, R., Riese, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viell, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (02-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 168426)

Birken, B., Nusbach, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukagater, B., Camarata, J., Chang, J.J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lu, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menes, I., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viell, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (28-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 28, 2003 this sequence version replaced gi:21702905.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L25102  
Center clone name: L45\_D\_9

----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 166432 bases at least Q40  
Consensus quality: 167350 bases at least Q30  
Consensus quality: 167611 bases at least Q20  
Insert size: 170000; agarose-  
Insert size: 167826; sum-of-contigs  
Quality coverage: 10.1 in Q20 bases; agarose-  
Quality coverage: 10.3 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
This sequence will be replaced  
By the finished sequence as soon as it is available and the accession number will be preserved.  
The accession number will be 1747 bp in length  
1748 1847: gap of 100 bp  
1848 103832: contig of 101985 bp in length  
103833 103932: gap of 100 bp

\* 103933 107966: contig of 4034 bp in length  
\* 107967 108066: gap of 100 bp  
\* 108067 119861: contig of 11795 bp in length  
\* 119862 119961: gap of 100 bp  
\* 119962 144243: contig of 24282 bp in length  
\* 144244 144343: gap of 100 bp  
\* 144344 165641: contig of 21298 bp in length  
\* 165642 185741: gap of 100 bp  
\* 185742 186426: contig of 2685 bp in length.

#### FEATURES

source  
1. 168426  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone\_lib="RP24-145D9"  
/clone\_lib="RP24-145D9"  
1. 1747  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
1848..103832  
/note="assembly\_fragment"  
103933..107966  
/note="assembly\_fragment"  
108067..119861  
/note="assembly\_fragment"  
119962..144243  
/note="assembly\_fragment"  
144344..165641  
/note="assembly\_fragment"  
165742..168426  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right

misc\_feature  
50157 a 33528 c 34710 g 49430 t 601 others  
BASE COUNT  
ORIGIN

Query Match 67.9%; Score 22.4; DB 2; Length 168426;  
Best Local Similarity 81.2%; Pred. No. 28;  
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GCTTAGACCACCATGGAGAGAGCGGAGATC 32  
DB 36738 GCTAGAGACCACCATGGAGAGATGCTACATC 36769

RESULT 4  
AC127560/c AC127560 186768 bp DNA linear ROD 04-JUN-2003  
DEFINITION Mus musculus chromosome 5 clone RP24-267118, complete sequence.  
ACCESSION AC127560  
VERSION AC127560.4 GI:31376488  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 186768)  
AUTHORS Wilson, R.K.  
TITLE The sequence of Mus musculus clone  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 186768)  
AUTHORS McPherson, J.D. and Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (17-JUL-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 3 (bases 1 to 186768)  
AUTHORS McPherson, J.D. and Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (08-MAR-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 4 (bases 1 to 186768)  
AUTHORS Wilson, R.K.



**TITLE** Direct Submission  
**JOURNAL** Submitted (04-JUN-2003) Genome Sequencing Center, 4444 For  
Parkway, St. Louis, MO 63108, USA  
**COMMENT** On Jun 4, 2003 this sequence version replaced gi:28867181.

```

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information
Center project name: M BB0267118
-----

```

FEATURES	source	Location/Qualifiers
		1..186768
		/organism="Mus musculus"
		/mol_type="genomic DNA"
		/db_xref="taxon:10090"
		/chromosome="5"
		/clone="RP24-367I18"
BASE COUNT	54278 a	37723 c 36959 g 57808 t
ORIGIN		

Sequence updated (06-Jun-2000).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 61 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1 11371: contig of 11371 bp in length
* 11372: gap of 100 bp
* 11472: contig of 10528 bp in length
* 22000: contig of 100 bp
* 22100: contig of 8478 bp in length
* 30578: gap of 100 bp
* 30578: contig of 100 bp
* 30578: contig of 6313 bp in length
* 36991: contig of 100 bp
* 37091: contig of 6230 bp in length
* 43321: contig of 100 bp
* 43321: contig of 5798 bp in length
* 49219: contig of 100 bp
* 49219: contig of 5270 bp in length
* 54589: contig of 100 bp
* 54589: contig of 5345 bp in length
* 60034: contig of 100 bp
* 60034: contig of 5916 bp in length
* 60334: contig of 100 bp
* 60334: contig of 5876 bp in length
* 66050: contig of 100 bp
* 66050: contig of 5876 bp in length
* 72026: contig of 100 bp
* 72026: contig of 4116 bp in length
* 72126: contig of 100 bp
* 72126: contig of 5426 bp in length
* 76242: contig of 100 bp
* 76242: contig of 100 bp
* 81768: contig of 100 bp
* 81768: contig of 4425 bp in length
* 86393: contig of 100 bp
* 86393: contig of 3142 bp in length
* 89335: contig of 100 bp
* 89335: contig of 3490 bp in length
* 93125: contig of 100 bp
* 93125: contig of 3539 bp in length
* 96764: contig of 100 bp
* 96764: contig of 4972 bp in length
* 101835: contig of 100 bp
* 101835: contig of 4219 bp in length
* 106154: contig of 100 bp
* 106154: contig of 4437 bp in length
* 106255: contig of 100 bp
* 106255: contig of 2753 bp in length
* 110592: contig of 100 bp
* 110592: contig of 100 bp
* 113544: contig of 100 bp
* 113544: contig of 3175 bp in length
* 116820: contig of 100 bp
* 116820: contig of 2351 bp in length
* 119270: contig of 100 bp
* 119270: contig of 3543 bp in length
* 122914: contig of 100 bp
* 122914: contig of 3139 bp in length
* 126153: contig of 100 bp
* 126153: contig of 3380 bp in length
* 129632: contig of 100 bp
* 129632: contig of 2822 bp in length
* 132554: contig of 100 bp
* 132554: contig of 2948 bp in length
* 135602: contig of 100 bp
* 135602: contig of 2123 bp in length
* 137023: contig of 100 bp
* 137023: contig of 100 bp
* 137826: contig of 100 bp
* 137826: contig of 2909 bp in length
* 140835: contig of 100 bp
* 140835: contig of 1354 bp in length
* 142289: contig of 100 bp
* 142289: contig of 3332 bp in length
* 145721: contig of 100 bp
* 145721: contig of 2480 bp in length
* 148301: contig of 100 bp

```

```

* 148401 150207: contig of 1807 bp in length
* 150208 150307: gap of 100 bp
* 150308 151625: contig of 1318 bp in length
* 151626 151725: gap of 100 bp
* 151726 153980: contig of 2255 bp in length
* 153981 154080: gap of 100 bp
* 154081 155858: contig of 1778 bp in length
* 155859 155958: gap of 100 bp
* 155959 157241: contig of 1283 bp in length
* 157242 157341: gap of 100 bp
* 157342 159107: contig of 1766 bp in length
* 159108 159207: gap of 100 bp
* 159208 160858: contig of 1651 bp in length
* 160859 160958: gap of 100 bp
* 160959 162417: contig of 1459 bp in length

Query Match 67.9%; Score 22.4; DB 2; Length 195637;
Best Local Similarity 81.2%; Pred. No. 28;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GCTTAGACCATGGGAGAGGCGAAGTC 32
Db 141850 GCTTAGACCATGGGAGAGGCGAAGTC 141881

RESULT 6
AC009052/c
LOCUS AC009052 200557 bp DNA linear PRI 19-MAR-2003
DEFINITION Homo sapiens chromosome 16 clone RP11-250E14, complete sequence.
ACCESSION AC009052
VERSION AC009052.8 GI:29124041
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 200557)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
DIRECT SUBMISSION
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 200557)
AUTHORS DOE Joint Genome Institute.
DIRECT SUBMISSION
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 200557)
AUTHORS DOE Joint Genome Institute.
DIRECT SUBMISSION
JOURNAL Submitted (06-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 200557)
AUTHORS DOE Joint Genome Institute.
DIRECT SUBMISSION
JOURNAL Submitted (01-MAY-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 5 (bases 1 to 200557)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
DIRECT SUBMISSION
JOURNAL Submitted (19-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Mar 19, 2003 this sequence version replaced gi:20376956.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www.sbgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.6.
LOCATION/Qualifiers
1. .200557
/organism="Homo sapiens"

FEATURES
source

```

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clones="RP11-250E14"
BASE COUNT 50926 a 47625 c 48640 g 53366 t
ORIGIN

Query Match 67.9%; Score 22.4; DB 9; Length 200557;
Best Local Similarity 81.2%; Pred. No. 28;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GCTTAGACACCATGGAGCAAGCGCAAGTC 32
Db 31590 GCTTAGACACCATGGAGCAAGCGCAAGCC 31559

RESULT 7
AC136859/c
LOCUS AC136859 203240 bp DNA linear HTG 23-NOV-2002
DEFINITION Rattus norvegicus clone CH230-482E23, WORKING DRAFT SEQUENCE.
ACCESSION AC136859
VERSION AC136859.2 GI:25188258
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Eureleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 203240)
Muzny,D.,Marle., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsebrook,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalio,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Deigado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gueorgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,D., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mahoney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokeleneh,O., Okwou,G., Olarnpusagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plapper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soza,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

```

Williams,G., Willson,R., Wlecyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

Unpublished  
Direct Submission  
2 (bases 1 to 203240)  
Rat Genome Sequencing Consortium.  
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 203240)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 23, 2002 this sequence version replaced gi:24818546.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GXJJ  
Center clone name: CH230-482E23  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 180755 bases at least Q40  
Consensus quality: 182011 bases at least Q30  
Consensus quality: 182628 bases at least Q20  
Estimated insert size: 186287; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as runs of N. The order of the pieces of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* By the finished sequence as soon as it is available and the accession number will be preserved.  
\* 1 203240: contig of 203240 bp in length.  
Location/Qualifiers  
1 .203240  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clones="CH230-482E23"  
1 .3086  
/note="wgs end extension  
clone end:Sp6"  
4114..5001  
/note="clone boundary  
clone end:Sp6  
site:Mbol

FEATURES  
source  
misc\_feature  
misc\_feature

```

misc_feature      end sequence:B2133954"
complement(200361..201269)
/nc="clone_boundary
clone_end:T7
site:Mbol
end sequence:B2133953"
BASE COUNT      49878 a 42514 c 41419 g 49473 t 19956 others
ORIGIN

Query Match      67.9%; Score 22.4; DB 2; Length 203240;
Best Local Similarity 81.2%; Pred. No. 28;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GCTCTAGACCACCATGGGAGGAGGCGAGTC 32
Db 89923 GCTCTAGAGCGATGGGAGGAGGCGAGTC 89892

RESULT 8
AC119491/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-498C19, WORKING DRAFT SEQUENCE, 2
unordered pieces.
AC119491
AC119491.6 GI:251337920
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 226999)
Muzny,D,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Ayalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,Y., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensukewa,L., Loulseghe,H., Lozano,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,B.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelimeh,O., Okwuonu,G., Olarnpunaagoo,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,
Flopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,D.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,

```

```

Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 226999)
Worley,K.C.
Direct Submission
Submitted (27-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 226999)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:22856259.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GK0J
Center clone name: CH230-498C19
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 194074 bases at least Q40
Consensus quality: 194867 bases at least Q30
Consensus quality: 195288 bases at least Q20
Estimated insert size: 204159; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 217820: contig of 217820 bp in length
* 217821 217920: gap of unknown length
* 217921 226999: contig of 9079 bp in length.
----- Location/Qualifiers
1. 226999
/organism="Rattus norvegicus"
/mol type="genomic DNA"
/db xref="taxon:10116"
/clone="CH230-498C19"
1. 1270
/notes="wgs contig"
BASE COUNT 54135 a 45923 c 44073 g 51585 t 31283 others

```

AUTHORS

Muzny, D., Marie, Metzker, M., Lee, Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, C., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, J., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyie, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, I., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Guevaratne, P., Geer, K., Gill, R., Grady, M., Guerra, W., Guevata, M., Gunaratne, P., Haaland W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., L., Hodgson, A., Hogues, J.M., Hollins, B., Howells, S., Rulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karkathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopes, J., Lorensuhewa, L., Loulsegod, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Minz, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidas, M., Murphy, M., Nair, D., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaekelem, O., Okwuonu, G., Olarunpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Pua, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmari, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, D., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

COMMENT

Direct Submission  
Unpublished  
2 (bases 1 to 239681)  
Worley, K.C.  
Direct Submission  
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 239681)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Nov 20, 2002 this sequence version replaced gi:3264889.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: KAYE  
Center clone name: CH230-98N15  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 229827 bases at least Q40  
Consensus quality: 232078 bases at least Q30  
Consensus quality: 233452 bases at least Q20  
Estimated insert size: 237409; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

**Misc feature**

```

1 .239681: 339681 bp in length.
Location/Qualifiers
1 . .239681
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-98N15"
1 . .1080
/note="wgs_end_extension
clone_end:17"
9276 . .9372
/note="clone_boundary
clone_end:17
site:
end_sequence:BH311747"
complement(237496 .238403)
/note="clone_boundary
clone_end:Sp6
site:"

```

[illegible]

/codon\_start=1  
/product="zinc finger protein 1"  
/protein\_id="AAA98913.1"  
/db\_xref="GI:1297186"  
/translation="MEPSIKGDEMUKKIKQHQDLELGLTLLSRGTATSSSELNLNLS  
KYTSSTSHQHQEQELADPRVFSNQCQKPYSSQALGGHQAHKERTIAKRGQY  
FKTSLPSFSAFAGSRYFASMASLPLGFSVNNRSTLGIQAHSITKPSFLRGQ  
TSLSHVFKQSIHQKPTIGKMLPEKPLEVAGNNNSMVAALKERIGHGFKSQNHQ  
FKKIDLTKL"  
complement(join(20081..20645,20700..22687))  
/note="Theoretical protein with similarity to protein  
encoded by GenBank Accession Number U41815, nucleoporin  
98"  
/codon\_start=1  
/protein\_id="AAA98914.1"  
/db\_xref="GI:1297187"  
/translation="MSTEISRMARDLDSKRRKRLSDGAAALCEHSEKIIDSLPMLNS  
PDYFLKPCINELVERETESDYSRVPDFTIGRIGYIRFLGNTDVRRLDLDHIVK  
HRHEIVDYDESGKPVVEGLKAAEVLTVNIPDLTWGQQVNHAYLKLKOSTERQ  
ATFISPDNDGLMKFFVPHSPRGLSDDEADMDADPGLDVPGLDGKQVADIDEE  
DOMETSLELSHSLPAHLGLDPEKMKEMLMFNEDESEDFRQTSMLTSLTKR  
NVAPSKIAQRSHQDPPVVRTPPLALDYNFNGDKSPGSLMQQKNLAVRKSK  
TGGFELDSHTVLTUNSNVVDAAFLMFRSGFRAGNGVLFHTGKPCISSQMV  
LSSVINKEKLAIDVWDRKGVKQKELIDSAFEAPLSLHKELNHVEEVRGFSFLK  
QNVTVDRVLSDCRSYIGIEKQLEVALGSTAKLFLKHQVWVVELIKVLFSEKST  
ERLWYAADNEEDVMQDVEDSAKIDTEALPLIRAEFSQWQESVSHVQVSDLN  
GSSYLEHLFLLTGRELDSAVELAKSGDVRACLISQAGGSTVNRNDILQHLWRR  
NGDFNPIEKERIKYELLAGNHLDLQDFTIDKRWPLGSLQLLNOAKAPWPIYI  
DEGPADGFDNDYHSDILYMLLHKEEDEFGLQTFMSAFSTDDPDLDHMIWHR  
GILEAVGAFSDMLDLMGFAQLLSQGLCHWAIYVHLHPREDPHVLRVTVIREI  
LFQVCEATSSMSQROPIKDLGPISEWMEALVTPYHSSFLLLLKLVLAHC"  
join(23732..23788,24275..24361,24477..24554,24641..24834,  
24949..25090,25275..25355,25408..25441,25883..25965)  
/note="Theoretical protein with similarity to Swiss-Prot  
Accession Number P41838 poly A+ RNA export protein"  
/codon\_start=1  
/protein\_id="AAA98915.1"  
/db\_xref="GI:1297188"  
/translation="MATFGPATANPNKNSYEVTPSPADSSISLFSRPRADLVATS  
WDNQVCWEISRGASLASAPKASISHDQPVLCWAKDDGTTFVSGGCDQAKMFWLL  
SQGPVTVAHEGPIVAMAMIPGMNLLATGSLQDPTLKYWDTRQCNVPHVQQLPKCYT  
LSVKHPLWVGTADRLNVLNLPONPOTEPKRIOSPLKYQTRCVTAFPDQDQGLELACS  
LDAYILDGAFNFDKSKQELVFSSCIDFLFC"  
join(27691..27759,27827..27964,28035..28919,28997..29101,  
29195..29504,29582..30096,30174..30347,30422..30604,  
30686..30886)  
/note="Theoretical protein with similarity to Swiss-Prot  
Accession Number P19456 plasma membrane Atpase 2 (proton  
pump)"  
/codon\_start=1  
/protein\_id="AAA98916.1"  
/db\_xref="GI:1297189"  
/translation="MKNPLSVNVELAAIYALANGGRPPDQDFVGLTVLLIINST  
ISITEENAGNAAALMAGLAPKTVLRDQKWEQEAAILVFGDIISIKLGIIVADG  
RLLDGDLKQALGELSPFTKHPQEVYSTCKQGEAAVAVIATGVHFFGKAA  
HLVDSTQEGHOKVTAIGNFCISIAIGMLIEIVYPIQKRAVDGIDNLLVLI  
GGIPAMPTVLSVTAIGSHLSQGAITKRWTAIEEMAGMDVLCSDIKTGLNKLT  
VDSKVEVFKLDKQLLVNAAARSRVENQDAIDACIVGMLGDPEAREEGTEVHF  
PRNPVDEKATITIDANGWHRYSKAPQIIELCNLRDASKAHDIDKADRLR  
SLAVQRTSEKDNKXGEPWQFLGLPLDPPRHSATIRKALDGVNVMITGDQ  
LAIGTQRELGKNTNPSYALLGQDQESIASLFDLIEKADGAFVLEHREYI  
VKLQEMKHTCGMTGQNDAPALRADIGIADATDAARSADIVLEAGLSVTS  
AVLTSAIRAFQKNYTIYASITIRVGMFLMALIKWDFSPFMVLIVAILNDGTIM  
TISKDRVKPSPLSKLEIPATGVGLTVLAVMTVFVFAESDFEFSAKGVRSI  
SGNPHELTAAYILOVSIQSALIPVTRSSWSVVERPSFWLISAFMAQLIATLAVY  
ANNFAKIGIGWAGVIVISYIFVLPDLILKFIIRVSLSGRADWNVIENKTAFS  
KDYGKGEREAQWAQRTLHGLQPAQTSDFMFDKSTYRELSEIADQAKRAEVARQ  
SSN"  
complement(join(33038..33163,33274..33363,33443..33691,  
34403..34644,34725..34895,35135..35297))  
/note="Theoretical protein with similarity to GenBank  
Accession Number L22302 serine/threonine protein kinase"  
/codon\_start=1

/protein\_id="AAA98917.1"  
/db\_xref="GI:1297190"  
/translation="WCLRLHLLPIWLLGPEFFSTANKSQSSVVYVSGASPEIPISQ  
ASPRNGAOSPPIVKVVRQLKNNKILIALIVSSLLCVTWFLVILLWRYNMKN  
SPTGIKRSDSVSVTKPTVHKIDSVRKGTIPVYEQLLSATAKNFSDSKVLRGGR  
GCLYRACLDKRSVTKKLDGGSTDEKFNKNLHKASEDLIDGKELIFFRGSIL  
TYMNEMLKPRFPCNHLGSKVTQNDVSVFGVILLELLGKRSVEKSSSESTVWA  
VPKLSDRANLNLIDPAIKTMDLKHLYQVAAVLVCQPEPSYRPLITDVLHSLPL  
LPVELGGSRLIL"  
BASE COUNT 11769 a 7092 c 7104 g 11602 t 3 others  
ORIGIN  
Query Match 66.7%; Score 22; DB 8; Length 37570;  
Best Local Similarity 83.3%; Pred. No. 49;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 2 CTCTAGACCACCATGGGAAGGCGAAGT 31  
|||||  
Db 10189 CTCTGTCACCATAGGAAGGGTGAAGT 10218  
|||||  
F23A5 109694 bp DNA linear PLN 02-DEC-1999  
Arabidopsis thaliana chromosome 1 BAC F23A5 sequence, complete  
sequence.  
AC011713  
AC011713.2 GI:6136350  
HTG.  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis;  
1 (bases 1 to 109694)  
Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S., Lee J.M., Li J.,  
Gonzalez A., Liu A., Liu K., Sakano H., Koo T., Pham P.,  
Vaysberg M., Hong B., Chin C., Choi E., Chou J., Altafi H.,  
Brooks S., Chao Q., Conn L., Conway A., Hansen N.,  
Johnson-Hopson C., Khan S., Kim C., Lam B., Nguyen M., Palm C.,  
Shinn P., Tambunga G., Davis R.W., Ecker J.R., Federepiel N.A. and  
Theologis A.  
Arabidopsis thaliana chromosome 1 BAC F23A5 sequence  
Unpublished  
REFERENCE  
2 (bases 1 to 109694)  
Theologis A.  
Direct Submission  
Submitted (13-OCT-1999) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
3 (bases 1 to 109694)  
Theologis A.  
Direct Submission  
Submitted (27-OCT-1999) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
4 (bases 1 to 109694)  
Theologis A.  
Direct Submission  
Submitted (02-DEC-1999) Plant Gene Expression Center, 800 Buchanan  
St., Albany, CA 94710, USA  
On Oct 27, 1999 this sequence version replaced gi:6041759.  
The sequence of BAC F23A5 from Arabidopsis thaliana chromosome 1.  
LOCATION/Qualifiers  
1. 109694  
/organism="Arabidopsis thaliana"  
/molecule\_type="genomic DNA"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="F23A5"  
complement(210..3931)  
/gene="F23A5.1"  
complement(join(210..306,392..573,655..837,912..1085,  
1163..1677,1755..2064,2158..2262,2340..3224,3294..3431,

```

3499. .3597,3681. .3800,3875. .3931))
/genes="F23A5.1"
/notes="Identical to gb|X737676 aha9 (ATAHA9) ATPase gene
from Arabidopsis thaliana."
/codon_start=1
/evidence=not_experimental
/protein_id="AAF14653.1"
/db_xref="GI:6503277"
/translation="WAGNKDSWDDIKNEGIDLEKIPBEVLITOLRCTREGITSDGQ
TRLEIFGPNKLESEKENKVLKPLGFWRNPLSWWELAAIMAILANGSGRPDPQDFV
GITVLLINSTISFIEENNAGNAALWAGLAPKTVLRDGRWSQEAAIIVPGDIIS
IKGLIVADPGDPLKIDQSALTGEISLPVTKPGQEVYSGTCKQLEAVVIA
TGVHFFGAALHSTOEBEHFQVLAIGNFCISAIAGMIEIIVMWDIQKRAV
DGLDNLVLLIGIPIAMPTVLVTMAIGSHLSOOGALTKEMTAIEVWAGIQLCD
KTGTLLNLTVDKSNVEVFDLDDOLLUNARASVENQDADACIAGVMDGPRE
AREGITEVHFFPNVDRKTAITYDANGNWHRSKGAPEQIIELCNLRBASRAHD
IIDKADRLSLAVGRQTVSEKDNKSGPGEQWGLLPLEDPPRHDSAEITRALDL
GVNVMITQDQAIQKETRGMLGTNMVPSALLGQDKESIASLPVDEIIEADGF
AGVFPHEKYEIVKRLQEMKHCIMGTGDVNDAPALGRADIGIADADTAARSADIV
LTEPGLSVISAVLSRAIFORMKNYTIYVNSITIRIVMGEMLLALIWKFDPSPFVWL
IVALLNDGIMTISDRVKPSPPLPSWKLKEIPATGVVLGTYLAWTVVFPWAESTD
AQLIATLIAYANWMPARIGIOWAGWAGWYIYVPIPLDILKFIIRYLSGRAND
NVIENKTATSKYDQKGERAQWQAQRTILHGLQPAQTSDFMENDKSTYSELSEIADQ
AKRAEVARLRERHTLKGHVSVKQKGLDIEAIQOHVTL"
complement(4792. .7526)
/genes="F23A5.2"
complement(join(4792. .4842,4939. .5019,5179. .5250,
5329. .5406,5512. .5640,5903. .5983,6168. .6309,6424. .6617,
6704. .6781,6897. .6983,7470. .7526))
/genes="F23A5.2"
/notes="Identical to gb|Q38942 RAE-like protein and
contains 5 PF|00400 WD domain, G-beta repeat domains.
ESTs gb|T76710, gb|Z37676, g"
/codon_start=1
/evidence=not_experimental
/product="F23A5.2(form2)"
/protein_id="AAF14655.1"
/db_xref="GI:6503279"
/translation="WATFGAPATANSNPNKSYEVTPSPADSISSFSFPRADILVATS
WNQVCWISRSGLASAPKASISHDQVLCSAWKDDGTVTFSGGCDKAKWPLL
SGQVTVVAHGGPIAAMAWIPQMLLATGSDWKTLKYWDTRQNPVHTQOLPKCYT
LSVQHPVLVGTADNLIVFNLQNPQTEFKRICSLKYQTRCVTAFPDQOQFLVGSIE
GRGVVHLDDSQOSKNFTKCRDNDIYVNSLNFHPVHGTFTAGSDGAFNWDKD
SQRLKAMSRGNQIPCCSFNHDGSIYAAACYDWSKGAENHNPNATAKSIIFLHPQE
SEYKAPRGATGRK"
complement(join(5293. .5375,5817. .5850,5903. .5983,
6168. .6309,6424. .6617,6704. .6781,6897. .6983,7470. .7526))
/genes="F23A5.2"
/notes="Identical to gb|Q38942 RAE-like protein and
contains 5 PF|00400 WD domain, G-beta repeat domains.
ESTs gb|T76710, gb|Z37676 an"
/codon_start=1
/evidence=not_experimental
/product="F23A5.2(form1)"
/protein_id="AAF14654.1"
/db_xref="GI:6503278"
/translation="WATFGAPATANSNPNKSYEVTPSPADSISSFSFPRADILVATS
WNQVCWISRSGLASAPKASISHDQVLCSAWKDDGTVTFSGGCDKAKWPLL
SGQVTVVAHGGPIAAMAWIPQMLLATGSDWKTLKYWDTRQNPVHTQOLPKCYT
LSVQHPVLVGTADNLIVFNLQNPQTEFKRICSLKYQTRCVTAFPDQOQFLVGSIE
GRGVVHLDDSQOSKNFTKCRDNDIYVNSLNFHPVHGTFTAGSDGAFNWDKD
SQRLKAMSRGNQIPCCSFNHDGSIYAAACYDWSKGAENHNPNATAKSIIFLHPQE
SEYKAPRGATGRK"
complement(5293. .5375,5817. .5850,5903. .5983,
6168. .6309,6424. .6617,6704. .6781,6897. .6983,7470. .7526))
/genes="F23A5.3"
/notes="Identical to gi|1297187 nucleoprotein 98 homolog
from Arabidopsis thaliana cosmid gb|U53501. EST gb|W43800
comes from this gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAF14656.1"
/db_xref="GI:6503280"

```

```

/translation="MSTEISRMARDUDSRKKRIRISLDGIAALCEHSKEIIDLPLMINS
PDYFLPCINELVERIEISPDYCSRPDFTIGRIGYGRIFLGTDTVRLLDDHIVKF
HHEVTVDDDESKPVVGGANKAAEVTAVNIPLDTWQGOQVNHVAYLKKOSTERQO
AFETSDPONGLWKFPVPHSPFGLSDDEADINDADGPGDGLDCKKVVADIDEE
DQMTSELSELSHLPAHLGLDPEKEMEMFNDEDEDEDFRQTSLSLMTSLTKR
NVPKQIAQRNSHQDPFPVVRVDTPLALFVNPNGNDKSPGSLVQVQKNLAVRSK
TGGFELDISHTVPLTDNYSRNVVDAFLMGRSGFRAGPNGVLFHTGKICSSSQMV
LSSVINKEKLIADKVVDKRGKVKOKELIDSAFEAPLSIKELNHVVEEVFGFSLKL
QSVVTRVLSIDICRVYIGIJEKQLEVAGLSAKILFLOHVMVVELIKVLPSEOST
ERLMAASNEEDVMQDVDESAKIDTEALPIIRAEFCWLOESVSHVQEDVSLN
GSSYDLHLFLTLGRDLDSAVELASIKGDRVJACILSQAGSGTNNRNDILQQHLWR
NGDFDFIKERIKUYELLAGNIHDALDFTIDWKRFLGSYQLLNQAKAPVPVYIY
DEGAPDGFVSDNKHSDILYLLMLLSKBEFEFGFLQTFMPSAFSTDDPDYHIMWHR
GILEAVGARTSDDLHTLDMGVFAQLLSQGLCHWAIYVVLHIIPREBPHVLTHTVIREI
LFQYCTWASMSORQFIKDLGIPBSMHEALVRYTHSHSFFLLLLKVLVAHC"
complement(13266. .14507)
/genes="F23A5.4"
complement(join(13266. .13528,13864. .14163,14375. .14507))
/genes="F23A5.4"
/notes="Contains similarity to gb|AF151904 CGI-146 protein
from Homo sapiens. EST gb|T4446 comes from this gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAF14657.1"
/db_xref="GI:6503281"
/translation="MLCRKNSLVDRGVVPLVNVYDLTPINGYAVMLGLGVVHSGVE
VHGIEYAYGAHEYPSTGIEGEPKQCEGTFKRSILIGKTDLGPLEVATMQLADNY
KGSYNIYITKNCHNFCDETCIKLTGNPISFWNRRLARIKFGMCMCVLPATINATR
FGNRYNMQDKSCBAENEKRLTSVSRERSTATIPSSSSSPSVQVGRSKRRRPAL
QPSSPLTLGSSSV"
15961. .17216
/genes="F23A5.5"
join(15961. .16202,16585. .16634,16729. .16781,16917. .17216)
/genes="F23A5.5"
/codon_start=1
/evidence=not_experimental
/product="F23A5.5"
/protein_id="AAF14659.1"
/db_xref="GI:6503283"
/translation="MIRGRGWSGLSTRUKLYTRRFTCKGENKLEKTSVAHHSDT
ESSVSYDETYKKLKLDFVTAAILTFEPKKNKFGDFVHVYQVFTIVCLPSVAVLYL
AQVARKMTDAELGCKRKEKEKEAOKALQVEAATKSHBELMEMRQLGKI
EETIKELVETKPKSGNAPTKQDQOSTKLSPEVQSPEKEHKGNVQKPGDGH"
complement(17482. .19779)
/genes="F23A5.6"
complement(join(17482. .17695,17927. .18022,18433. .18839,
18946. .19779))
/genes="F23A5.6"
/notes="Contains similarity to gb|AF092102 G-protein beta
subunit gitsp from Schizosaccharomyces pombe and contains
2 PF|00400WD domain, G-beta repeat domains."
/codon_start=1
/evidence=not_experimental
/protein_id="AAF14658.1"
/db_xref="GI:6503282"
/translation="MATEYERKLENIRNDEMLAALNVRKASSLLSAKR3RDSK
SLKAPTEKPASTEYIRMSLRITGLNPDAGLPDGFDFMGSQITHNQSPQKQSPR
FFAPIPEGAEGYGYTQVLTDLGIESKRCRKLKGEIGVYKDNESPMVTRSS
SRVYKSVKKEBEDDSDFYVNKFEIPVKPEIFEDLDLLTLPQNVARVVPGRIF
VVOFLPCENVMVAGDKLGNVGFNLDLCSNEEDNDIGIYLFTHPSAFSVIVQONS
SRVSSYDGLIRLMDVEKSVPLVYVSTDEAFSLSQRENDQSILFGQDYGVTNWD
LRAGSVFWELHERINSIDFNQPNPHVATSSDTGTACLMDLRSGAKKPKLTSTV
NHSRAVSAFSPSGSLATSLDNYIGVLSGANFENTCMIYHNNTSRWISKFAYWG
WDDSYIVGNLSKKIDVINPKLRITVWELNPLQRAIPCRHCHPNVGTLAGSTAGG
QVTVWTTK"
20272. .21330
/genes="F23A5.7"
join(20272. .20358,20752. .21074,21168. .21330)
/genes="F23A5.7"
/codon_start=1
/evidence=not_experimental
/product="F23A5.7"
/protein_id="AAF14660.1"

```



/db\_xref="GI:6503284"  
/translation="MQSAFQGNILIRNEIEYELHSPPLQPPNSFGFTVDPRFG  
EQMISLRNFKDKEDIKTIFMTFSTKTEPEYILHITFIWISKAGATEAL  
EIMCSAMPDITIEISKLIRRGINTSPSYGPFEEELDDQLQALQVLEERGISDEL

Query Match 66.7%; Score 22; DB 8; Length 109694;  
Best Local Similarity 83.3%; Pred. No. 44;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 CTCTAGACCACCATGGGAGGAGGCGAGT 31  
Db 21069 CTCTGGTCCACCATAGGAGGCGTGAAGT 21040

RESULT 13  
AL671893/c  
LOCUS  
DEFINITION  
AL671893  
VERSION  
AL671893.7 GI:22759503  
KEYWORDS  
HTG.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 175057)  
Lawlor, S.  
Direct Submission  
Submitted (30-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk  
On Sep 9, 2002 this sequence version replaced gi:2265401.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
-----

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em: EMBL; Sw:  
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP  
database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-170G19 is  
from the RPCI-23 Mouse PAC Library  
constructed by the group of Pieter de Jong.  
For further details, see <http://www.chori.org/bacpac/home.htm>  
VECTOR: PBACe3.6.

FEATURES  
Location/Qualifiers  
1. 175057  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="X"  
/clone="RP23-170G19"  
/clone\_lib="RPCI-23"  
BASE COUNT 52683 a 37591 c 36564 g 48219 t  
ORIGIN

Query Match 66.7%; Score 22; DB 10; Length 175057;

Best Local Similarity 83.3%; Pred. No. 43;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 GCTCTAGACCACCATGGGAGGAGGCGAGT 30  
Db 128835 GCTCTAGAAACCATGGATAGTGGAGG 128806

RESULT 14  
AC120853  
LOCUS  
DEFINITION  
AC120853  
VERSION  
AC120853.3 GI:28394405  
KEYWORDS  
HTG; HTGS PHASE1; HTGS DRAFT.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 218976)  
Birren, B., Nusbaum, C. and Lander, E.  
Mus musculus chromosome 17, clone RP23-343A20  
Unpublished  
2 (bases 1 to 218976)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
Cook, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Farro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,  
Lander, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, M., McEwan, F., McKernan, K., Meldrum, J., Meneus, L.,  
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Olivier, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Triggilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 218976)  
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,  
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,  
Collymore, A., Cook, A., Cooke, P., Corum, B., DeAtellano, K.,  
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Farro, S.,  
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,  
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,  
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,  
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,  
Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,  
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,  
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,  
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (15-FEB-2003) Whitehead Institute/MIT Center for Genome

Consensus quality: 25/301 bases at least Q20  
-----

```
...
/note="assembly_fragment"
```

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 5576: contig of 5576 bp in length  
\* 5577 5676: gap of unknown length  
\* 5677 10157: contig of 4481 bp in length  
\* 10158 10257: gap of unknown length  
\* 10258 20834: contig of 10577 bp in length  
\* 20835 20935: gap of unknown length  
\* 20935 34257: contig of 13323 bp in length  
\* 34258 34357: gap of unknown length  
\* 34358 48319: contig of 13962 bp in length  
\* 48320 48419: gap of unknown length  
\* 48420 64650: contig of 16231 bp in length  
\* 64651 64750: gap of unknown length  
\* 64751 99005: contig of 34255 bp in length  
\* 99006 99105: gap of unknown length  
\* 99106 120795: contig of 21690 bp in length  
\* 120796 120895: gap of unknown length  
\* 120896 150392: contig of 29497 bp in length  
\* 150393 150492: gap of unknown length  
\* 150493 176013: contig of 25521 bp in length  
\* 176014 176114: gap of unknown length  
\* 176114 201237: contig of 25124 bp in length  
\* 201238 201337: gap of unknown length  
\* 201338 240155: contig of 38818 bp in length.

Location/Qualifiers

1. 240155  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="UNK"  
/clone="RP24-403B24"  
1. 5576  
/note="assembly\_name:Contig22"  
5677..10157  
/note="assembly\_name:Contig23"  
10258..20834  
/note="assembly\_name:Contig24"  
20935..34257  
/note="assembly\_name:Contig25"  
34358..48319  
/note="assembly\_name:Contig26"  
48420..64650  
/note="assembly\_name:Contig27"  
64751..99005  
/note="assembly\_name:Contig28"  
99106..120795  
/note="assembly\_name:Contig29"  
120896..150392  
/note="assembly\_name:Contig30"  
150493..176013  
/note="assembly\_name:Contig31"  
176114..201237  
/note="assembly\_name:Contig32"  
201338..240155  
/note="assembly\_name:Contig33"  
65110 a 54275 c 54006 g 65608 t 1156 others

BASE COUNT

Query Match 66.7%; Score 22; DB 2; Length 240155;

Best Local Similarity 83.3%; Pred. No. 41;

Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 4 CTAGCACCACCATGGGAGAGCGGAGTCA 33  
|||||  
Db 218613 CTAGCCACCATTTGGCAGAGCGGAGTCA 218584  
|||||

Search completed: December 3, 2003, 13:18:39  
Job time : 904.273 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:41 ; Search time 157.903 Seconds  
(without alignments)  
564.131 Million cell updates/sec

Title: US-10-082-772-1

Perfect score: 33

Sequence: 1 gcttagaccacatgggaagaaggcgaagtcga 33

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03:\*

1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*

2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*

3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*

4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*

5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*

6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*

7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*

8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*

9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*

10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*

11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*

12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*

13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*

14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*

15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*

16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*

17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*

18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*

19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*

20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*

21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*

22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*

23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*

24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

25: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	33	AAF61392	Expression vector
2	22	66.7	933	AAC48285	Arabidopsis thalia
3	22	66.7	935	AAC34608	Arabidopsis thalia
4	21	63.6	35	ACC44614	Lambda integrase p
5	21	63.6	5855	ACC44715	Plasmid pCXlanIntR
6	21	63.6	5855	ACC44730	Plasmid pCX-LamInt
7	21	63.6	5855	ABT16616	Artificial plant c
8	20.6	62.4	717	AAK69042	Human immune/haema

9	20.6	62.4	2244	21	AAZ92416	cDNA encoding huma
10	20.6	62.4	3198	22	AAH14574	Human cDNA sequenc
11	20.6	62.4	3760	22	AAH17715	Human cDNA sequenc
12	20.6	62.4	3941	22	RAAD03459	Human ubiquitin pr
13	20.6	62.4	4650	22	AAH17670	Human cDNA sequenc
14	20.4	61.8	5549	24	ABK91528	Modified HIV prote
15	20.4	61.8	5549	24	ABK91529	Modified HIV prote
16	20.4	61.8	5549	24	ABK91530	Modified HIV prote
17	20.4	61.8	5549	24	ABK91531	Modified HIV prote
18	20.6	60.6	55	21	AAK93056	Oligonucleotide #2
19	20.6	60.6	55	25	ABX11660	PCR primer #2 used
20	20.6	60.6	243	24	ABN93007	Staphylococcus epi
21	20.6	60.6	285	24	ABL77925	Human ovarian canc
22	20.6	60.6	331	24	ABL77939	Human ovarian canc
23	20.6	60.6	340	24	ABL77602	Human ovarian canc
24	20.6	60.6	1071	25	ACC44648	Lambda integrase E
25	20.6	60.6	1071	25	ABT16637	Artificial plant c
26	19.8	60.0	1890	22	AAH57559	Human brain cell s
27	19.6	59.4	6189	23	AAH59650	Propionibacterium
28	19.4	58.8	34	21	AAH15340	PCR primer for a N
29	19.4	58.8	747	22	AAH62738	Shrimp white spot
30	19.4	58.8	4087	22	AAH27712	DNA encoding novel
31	19.4	58.8	5244	22	AAH41189	Human oligonucleot
32	19.4	58.8	9448	22	AAH41187	Human oligonucleot
33	19.4	58.8	9771	22	AAK71552	Human immune/haema
34	19.4	58.8	14676	22	AAK71553	Human immune/haema
35	19.4	58.8	15496	22	AAK71549	Human immune/haema
36	19.4	58.8	18830	22	AAK71558	Human immune/haema
37	19.4	58.8	32217	22	AAH41742	Geromic sequence #
38	19.4	58.8	59560	22	AAK71550	Human immune/haema
39	19.4	58.8	59560	22	AAK80620	Human immune/haema
40	19.4	58.8	305107	22	AAH62689	Shrimp white spot
41	19.2	58.2	392	22	AAH03415	Human reproductive
42	19.2	58.2	392	22	AAH03417	Human reproductive
43	19.2	58.2	396	24	ABN61714	Human cancer relat
44	19.2	58.2	401	22	AAH03416	Human reproductive
45	19.2	58.2	401	22	AAK89237	Human digestive sy

ALIGNMENTS

RESULT 1

AAF61392

ID AAF61392 standard; DNA; 33 BP.

AC AAF61392;

XX

DT 05-JUN-2001 (first entry)

XX

DE Expression vector pKEX-2-XR Int gene PCR primer 3343.

DE

XX Sequence-specific recombination; SSR; integrase; gene therapy; somatic;

KW targeted integration; Int gene; PCR primer; ss.

KW

XX Synthetic.

OS

XX

PN DE19941186-A1.

PN

XX

PD 01-MAR-2001.

XX

PF 30-AUG-1999; 99DE-1041186.

XX

PR 30-AUG-1999; 99DE-1041186.

PR

FA (DROE/) DROEGE P.

FA

XX Droege P;

PI

XX WPI; 2001-246016/26.

DR

XX Sequence-specific recombination of DNA in eukaryotes, useful

PT particularly for somatic cell gene therapy, uses an integrase to effect

```

PT recombination between att sites -
XX Example 1; Page 8; 24pp; German.
PS
XX
CC This invention describes a novel sequence-specific recombination (SSR) of
CC DNA in a eukaryotic cell, comprising introducing two DNA sequences (I,
CC II) into a cell, using an integrase (Int) to effect SSR. The invention
CC also describes (1) a nucleic acid comprising a 243 base pair sequence
CC (III), fully defined in the specification, or its derivatives; and (2)
CC a vector containing (III), or its derivatives, plus a therapeutic gene, or
CC its derivatives. The method is particularly used in somatic gene therapy
CC in humans and animals, but can be applied more generally for gene
CC transfer to animal or plant cells. The method is simple and controllable,
CC and provides stable and targeted integration of selected DNA sequences.
XX
SQ Sequence 33 BP; 11 A; 8 C; 10 G; 4 T; 0 other;

Query Match 100.0%; Score 33; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTCTAGACACCATGGGAAGGCGGAAGTCA 33
Db 1 GCTCTAGACACCATGGGAAGGCGGAAGTCA 33

RESULT 2
AAC48285/c
ID AAC48285 standard; DNA; 933 BP.
XX
AC AAC48285;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 56921.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW Protein identification; signal transduction pathway;
KW Metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX
PN EP1033405-A2.
XX
PD
XX
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

```

```
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149369.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 66.7%; Score 22; DB 21; Length 933;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CTCTAGACCACCATGGGAGGAGGGAAGT 31
Db 554 CTCTGGTCCACCATAGGAGGAGGTTGAAGT 525

RESULT 3
AAC34608/c
ID AAC34608 standard; DNA; 935 BP.
XX AAC34608;
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7243.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
```

PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140921.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144321.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148664.  
PR 16-AUG-1999; 99US-0149358.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161320.  
PR 28-OCT-1999; 99US-0161992.







Db 1724 CATGGGAAGAGGCGAAGTCA 1744  
RESULT 8  
ID AAK69042/c  
AC AAK69042 standard; DNA; 717 BP.  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23854.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW Cystostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216680.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0248526.  
PR 08-NOV-2000; 2000US-0248527.  
PR 08-NOV-2000; 2000US-0248528.  
PR 08-NOV-2000; 2000US-0248532.  
PR 08-NOV-2000; 2000US-0248609.  
PR 08-NOV-2000; 2000US-0248610.  
PR 08-NOV-2000; 2000US-0248611.  
PR 08-NOV-2000; 2000US-0248613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.

```
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 03-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX Disclosure; SEQ ID NO 23854; 3071pp + Sequence Listing; English.
XX AA54951 to AA64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AA82170 to AA82171. (I) have cytosolic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AA64703
CC to AA87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AA54942 to AA54950 and AA82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 717 BP; 120 A; 197 C; 235 G; 165 T; 0 other;
Query Match 62.4%; Score 20.6; DB 22; Length 717;
Best Local Similarity 85.2%; Pred. No. 45;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 CTAGACCAACCATGGGAAGAGGCGAAG 30
Db 603 CTAGACGACCCCTGGAAGAAGAGGTGAAG 577
RESULT 9
AAZ92416
ID AAZ92416 standard; cDNA; 2244 BP.
XX
AC AAZ92416;
XX
XX 05-JUN-2000 (first entry)
DT
XX cDNA encoding human deubiquitinating protein Dub12.
DE
XX Human; deubiquitinating protein; Dub11; Dub12; immune disorder;
KW inflammation; allergy; immunosuppressant; antiarthritic; antirheumatoid;
KW antiinflammatory; dermatological; antithyroid; ss.
XX
OS Homo sapiens.
XX
```

```
FH Key Location/Qualifiers
FT CDS 1..2244
FT /partial
FT /*tag= a
FT /product= "Human Dub12"
FT note= "No stop codon given in the specification"
XX WO200001817-A2.
XX 13-JAN-2000.
XX 06-JUL-1999; 99WO-US12366.
XX 06-JUL-1998; 98US-0110938.
XX 13-JUL-1998; 98US-0114466.
XX 23-JUL-1998; 98US-0093897.
XX 12-AUG-1998; 98US-0132968.
XX 18-AUG-1998; 98US-0136214.
XX 11-SEP-1998; 98US-0099999.
XX (SCHE ) SCHERING CORP.
XX Bates EM, Lebecque SUE, Murphy EE, Mattson JD, Gorman DM;
XX Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
XX Bazan JF, Mahony D, Lees EM;
XX WPI; 2000-171015/15.
XX P-PSDB; AAY77473.
XX New isolated mammalian genes, used to develop products for treating
PT e.g. immune, inflammatory or allergic abnormalities, cancers or
PT degenerative conditions -
XX Claim 60; Page 192-195; 218pp; English.
XX The invention relates to a number of primate and/or rodent proteins, and
CC the genes which encode them. The invention encompasses human dendritic
CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
CC factor) receptor family-related proteins HTRA84, HSLJD37R and RANKL;
CC human CC chemokine HC65; human deubiquitinating proteins Dub11 and Dub
CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
CC properties of ligands for proteins comprising a leucine-rich motif
CC (LRR); human cyclin E2; cDNAs encoding these proteins; and antibodies
CC against these proteins. The proteins can be used for modulating the
CC physiology or development of a cell. They can be used to mediate uptake
CC of substrates (e.g., prostaglandin-like molecules), to modulate or
CC mediate cellular interactions (e.g., induce or prevent trafficking,
CC proliferation, or differentiation of cells), or are intracellular
CC proteins which are important in various cellular processes such as the
CC deubiquitination of proteins or cell cycle regulation. The products can
CC be used for treating medical conditions such as immune, inflammatory or
CC allergic disorders, or abnormal cellular proliferation, for example,
CC cancers or degenerative conditions. They can be used to modulate immune
CC responses in disease states e.g., autoimmune disorders, including
CC rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's
CC autoimmune thyroiditis, as well as acute and chronic inflammatory
CC responses in which T cell activation, expansion, and/or immunological T
CC cell memory play an important role. Sequences AAZ92413-Z92416 represent
CC cDNAs encoding human deubiquitinating proteins (Dub). AAZ92413-Z92414
CC encode Dub11 proteins and AAZ92415-Z92416 encode Dub12 proteins.
XX
SQ Sequence 2244 BP; 552 A; 714 C; 598 G; 380 T; 0 other;
Query Match 62.4%; Score 20.6; DB 21; Length 2244;
Best Local Similarity 85.2%; Pred. No. 52;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 CTAGACCAACCATGGGAAGAGGCGAAG 30
Db 928 CTAGACGACCCCTGGAAGAAGAGGTGAAG 954
RESULT 10
```

```
AAH14574
ID AAH14574 standard; cDNA; 3198 BP.
AC AAH14574;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12163.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 12163; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 3198 BP; 746 A; 1013 C; 881 G; 558 T; 0 other;
XX
Query Match 62.4%; Score 20.6; DB 22; Length 3198;
Best Local Similarity 85.2%; Pred. No. 55;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 CTAGACCCATGGGAAGAGCGGAAG 30
DB 1921 CTAGACCCCTGGAAAGAGGTGAAG 1947

AAH17715
ID AAH17715 standard; cDNA; 3760 BP.
AC AAH17715;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:17312.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 17312; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 3760 BP; 921 A; 1122 C; 1023 G; 694 T; 0 other;
XX
Query Match 62.4%; Score 20.6; DB 22; Length 3760;
Best Local Similarity 85.2%; Pred. No. 56;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 CTAGACCCATGGGAAGAGCGGAAG 30
DB 1828 CTAGACGACCTGGAAAGAGGTGAAG 1854
```

Best Local Similarity 85.2%; Pred. No. 56;  
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0

QY 4 CTAGACCACCATGGGAGAGAGCGGAAG 30  
|||||  
DB 1863 CTAGACGACCTGGGAGAGAGGTGAAG 1889

RESULT 13  
AAH17670  
ID AAH17670 standard; cDNA; 4650 BP.  
XX  
AC AAH17670;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human cDNA sequence SEQ ID NO:17234.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
FN BP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-0116126.  
XX  
PR 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
DR WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602  
full-length cDNAs defined in the specification, and for the detection  
of and/or diagnosis of the abnormality of the proteins encoded by the  
full-length cDNAs -  
XX  
PS Claim 8; SEQ ID 17234; 2537pp + CD ROM; English.  
XX  
CC The present invention describes primer sets for synthesising 5602  
full-length cDNAs defined in the specification. Where a primer set  
comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
to the complementary strand of a polynucleotide which comprises one of  
the 5602 nucleotide sequences defined in the specification, where the  
oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
of an oligonucleotide comprising a sequence complementary to the  
complementary strand of a polynucleotide which comprises a 5'-end  
sequence and an oligonucleotide comprising a sequence complementary to a  
polynucleotide which comprises a 3'-end sequence, where the  
oligonucleotide comprises at least 15 nucleotides and the combination of  
the 5'-end sequence/3'-end sequence is selected from those defined in  
the specification. The primer sets can be used in antisense therapy and  
in gene therapy. The primers are useful for synthesising polynucleotides,  
particularly full-length cDNAs. The primers are also useful for the  
detection and/or diagnosis of the abnormality of the proteins encoded by  
the full-length cDNAs. The primers allow obtaining of the full-length  
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
represent oligonucleotides, all of which are used in the exemplification  
of the present invention.  
XX  
SQ Sequence 4650 BP; 1045 A; 1263 C; 1355 G; 987 T; 0 other;

RESULT 12  
AAD03459  
ID AAD03459 standard; cDNA; 3941 BP.  
XX  
AC AAD03459;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
DE Human ubiquitin protease cDNA.  
XX  
KW Human; ubiquitin protease; cytostatic; virucidal; gene therapy; tumour;  
neoplasia; hepatic injury; cirrhosis; hepatitis; heart failure; rickets;  
myocardial infarction; skeletal muscle tumour; Grave's disease; atrophy;  
glomerulonephritis; cryptorchidism; periductal mastitis; Paget's disease;  
cystic renal dysplasia; polycystic ovarian disease; cretinism; myxoedema;  
inflammation; obstructive pulmonary disease; osteoporosis; enterocolitis;  
idiopathic inflammatory bowel disease; bronchial carcinoma; osteomalacia;  
choriocarcinoma; nodular hyperplasia; metastatic tumour; thyroiditis; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 279..3650  
FT FT /\*tag= a  
FT FT /\*product= "Human ubiquitin protease"  
XX  
PN WO200123585-A1.  
XX  
XX  
PD 05-APR-2001.  
XX  
PF 29-SEP-2000; 2000WO-US26915.  
XX  
PR 30-SEP-1999; 99US-0408865.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Kapeller-Libermann R;  
XX  
DR WPI; 2001-235374/24.  
DR P-PSDB; AAY71889.  
XX  
PT Nucleic acids encoding human ubiquitin protease polypeptides, useful  
for preventing, diagnosing and treating, e.g. breast, lung or liver  
cancer and viral infections -  
XX  
PS Claim 2; Fig 1; 120pp; English.  
XX  
CC The present cDNA sequence encodes human ubiquitin protease belonging to  
the family of mammalian deubiquitinating enzymes. This ubiquitin protease  
is encoded by the cDNA insert of plasmid deposited with ATCC NO:  
PRA-1849. Ubiquitin protease is highly expressed in foetal kidney,  
testes, foetal liver, ovary and foetal heart. It is also expressed in  
liver, breast, lung, and colon tissue and in liver metastasis derived  
from malignant colon tissues. Ubiquitin protease is useful for treating  
disorders mediated by or related to deubiquitinating enzymes, ubiquitin  
and ubiquitin protease-associated disorders, including neoplasia and  
tumours of the breast, lung, liver and viral infections. Ubiquitin  
protease is useful for treating disorders of liver (hepatic injury;  
cirrhosis, hepatitis); heart (heart failure, myocardial infarction);  
kidney (cystic renal dysplasia, glomerulonephritis); breast (periductal  
mastitis, Paget's disease); testis and epididymis (cryptorchidism,  
atrophy, inflammations); prostate (nodular hyperplasia); thyroid  
(thyroiditis, Grave's disease, cretinism, myxoedema); lung (bronchial  
carcinoma, obstructive pulmonary disease); colon (idiopathic inflammatory  
bowel disease, enterocolitis); ovary (polycystic ovarian disease,  
choriocarcinoma); bone disorder (osteoporosis, rickets, osteomalacia) and  
skeletal muscle tumours such as rhabdomyosarcoma. Ubiquitin protease DNA  
is also useful in gene therapy.  
XX  
SQ Sequence 3941 BP; 961 A; 1181 C; 1076 G; 723 T; 0 other;

Query Match 62.4%; Score 20.6; DB 22; Length 4650;  
Best Local Similarity 85.2%; Pred. No. 58;  
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 CTAGACACCATGGGAGAGGCGAAG 30  
DB 2141 CTAGACACCATGGGAGAGGCGAAG 2167

RESULT 14  
ABK91528  
ID ABK91528 standard; DNA; 5549 BP.

XX AC ABK91528;  
XX DT 14-AUG-2002 (first entry)  
XX DE Modified HIV protein-encoding plasmid DNA #80.

XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;  
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;  
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX Synthetic.  
XX OS  
XX PN WO200232943-A2.  
XX PD 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.  
XX PF  
XX PR 14-AUG-2000; 2000US-225097P.  
XX PR 14-NOV-2000; 2000US-252115P.  
XX PR 28-MAR-2001; 2001US-279257P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX (CHAD/) CHADRABARTI B K.

XX Nabel GJ, Huang Y;  
XX WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and  
PT its encoded protein, useful as vaccines for genetic or protein  
PT immunisation for acquired immunodeficiency syndrome or human  
PT immunodeficiency virus infection -  
XX Disclosure; Page 562-564; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV  
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its  
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for  
CC genetic or protein immunisation to a host, respectively. In particular  
CC these are useful for ameliorating the symptoms of acquired  
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,  
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
CC compositions are useful for treating or preventing HIV infections or  
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding  
CC plasmid DNA sequences of the invention.

SQ Sequence 5549 BP; 1362 A; 1471 C; 1393 G; 1323 T; 0 other;  
Query Match 61.8%; Score 20.4; DB 24; Length 5549;  
Best Local Similarity 80.0%; Pred. No. 72;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GCTCTAGACCATGGGAGAGGCGAAG 30  
DB 1903 GCTCTAGACCATGGGAGAGGCGAAG 1932

RESULT 15

ABK91529  
ID ABK91529 standard; DNA; 5549 BP.  
XX AC ABK91529;  
XX DT 14-AUG-2002 (first entry)  
XX DE Modified HIV protein-encoding plasmid DNA #81.

XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;  
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;  
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX Synthetic.  
XX OS  
XX PN WO200232943-A2.  
XX PD 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.  
XX PF  
XX PR 14-AUG-2000; 2000US-225097P.  
XX PR 14-NOV-2000; 2000US-252115P.  
XX PR 28-MAR-2001; 2001US-279257P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX (CHAD/) CHADRABARTI B K.

XX Nabel GJ, Huang Y;  
XX WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and  
PT its encoded protein, useful as vaccines for genetic or protein  
PT immunisation for acquired immunodeficiency syndrome or human  
PT immunodeficiency virus infection -  
XX Disclosure; Page 564-566; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV  
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its  
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for  
CC genetic or protein immunisation to a host, respectively. In particular  
CC these are useful for ameliorating the symptoms of acquired  
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,  
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
CC compositions are useful for treating or preventing HIV infections or  
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding  
CC plasmid DNA sequences of the invention.

SQ Sequence 5549 BP; 1366 A; 1463 C; 1397 G; 1323 T; 0 other;  
Query Match 61.8%; Score 20.4; DB 24; Length 5549;  
Best Local Similarity 80.0%; Pred. No. 72;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GCTCTAGACCATGGGAGAGGCGAAG 30  
DB 1903 GCTCTAGACCATGGGAGAGGCGAAG 1932

Search completed: December 3, 2003, 12:23:34  
Job time : 159.909 secs







OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/634,278  
FILING DATE: 19-DEC-1990  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 110:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 114 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (oligonucleotide)  
US-07-634-278-110

Query Match 57.6%; Score 19; DB 1; Length 114;  
Best Local Similarity 81.5%; Pred. No. 19;  
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 TCTAGACCACCATGGGAAGGCGAA 29  
Db 5 TCTAGACCACCATGGGATGGAGCTGGA 31

RESULT 6  
US-08-477-728-110  
Sequence 110, Application US/08477728  
Patent No. 5585089  
GENERAL INFORMATION:  
APPLICANT: QUEEN, Cary L.  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: SELICK, Harold E.  
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,728  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/634,278  
FILING DATE: 19-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 110:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 114 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (oligonucleotide)  
US-08-477-728-110

Query Match 57.6%; Score 19; DB 1; Length 114;  
Best Local Similarity 81.5%; Pred. No. 19;  
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 TCTAGACCACCATGGGAAGGCGAA 29  
Db 5 TCTAGACCACCATGGGATGGAGCTGGA 31

RESULT 7  
US-08-474-040-110  
Sequence 110, Application US/08474040  
Patent No. 5693761  
GENERAL INFORMATION:  
APPLICANT: QUEEN, Cary L.  
APPLICANT: CO, Man Sung  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: LANDOLFI, Nicholas F.  
APPLICANT: COELINGH, Kathleen L.  
APPLICANT: SELICK, Harold E.  
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Khourie and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,040  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/634,278  
FILING DATE: 19-DEC-1990  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223

```
/ REFERENCE/DOCKET NUMBER: 11823-002600
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 326-2400
/ TELEFAX: (415) 326-2422
/ INFORMATION FOR SEQ ID NO: 110:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 114 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (oligonucleotide)
/ US-08-474-040-110

Query Match          57.6%; Score 19; DB 1; Length 114;
Best Local Similarity 81.5%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 3 TCTAGACCACCATGGGAGGAGCGGAA 29
Db 5 TCTAGACCACCATGGGAGGAGCGTGA 31

RESULT 8
US-08-487-200-110
; Sequence 110, Application US/08487200
; Patent No. 5693762
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,200
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 19-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 base pairs
```

```
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (oligonucleotide)
/ US-08-487-200-110

Query Match          57.6%; Score 19; DB 1; Length 114;
Best Local Similarity 81.5%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 3 TCTAGACCACCATGGGAGGAGCGGAA 29
Db 5 TCTAGACCACCATGGGAGGAGCGTGA 31

RESULT 9
US-08-484-537-110
; Sequence 110, Application US/08484537
; Patent No. 6180370
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,537
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/634,278
; FILING DATE: 19-DEC-1990
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (oligonucleotide)
/ US-08-484-537-110

Query Match          57.6%; Score 19; DB 3; Length 114;
Best Local Similarity 81.5%; Pred. No. 19;
```

```
Matches 22; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 3 TCTAGACCACCATGGGAGAGCGGAA 29
    |||||
Db 5 TCTAGACCACCATGGGAGAGCGGAA 31

RESULT 10
US-07-634-278-74
; Sequence 74, Application US/07634278
; Patent No. 5530101
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/634,278
; FILING DATE: 19-DEC-1990
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (oligonucleotide)
US-07-634-278-74

Query Match 57.6%; Score 19; DB 1; Length 132;
Best Local Similarity 81.5%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 3 TCTAGACCACCATGGGAGAGCGGAA 29
    |||||
Db 5 TCTAGACCACCATGGGAGAGCGGAA 31

RESULT 11
US-08-477-728-74
; Sequence 74, Application US/08477728
```

```
; Patent No. 5585089
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th floor
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,728
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 19-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (oligonucleotide)
US-08-477-728-74

Query Match 57.6%; Score 19; DB 1; Length 132;
Best Local Similarity 81.5%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 3 TCTAGACCACCATGGGAGAGCGGAA 29
    |||||
Db 5 TCTAGACCACCATGGGAGAGCGGAA 31

RESULT 12
US-08-474-040-74
; Sequence 74, Application US/08474040
; Patent No. 5693761
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
```

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Townsend and Townsend Khourie and Crew  
;; STREET: 379 Lytton Avenue  
;; CITY: Palo Alto  
;; STATE: California  
;; COUNTRY: US  
;; ZIP: 94301  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/474,040  
;; FILING DATE: 07-JUN-1995  
;; CLASSIFICATION: 536  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/634,278  
;; FILING DATE: 19-DEC-1990  
;; APPLICATION NUMBER: US 07/590,274  
;; FILING DATE: 28-SEP-1990  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/310,252  
;; FILING DATE: 13-FEB-1989  
;; APPLICATION NUMBER: US 07/290,975  
;; FILING DATE: 28-DEC-1988  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Smith, William M  
;; REGISTRATION NUMBER: 30,223  
;; REFERENCE/DOCKET NUMBER: 11823-002600  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 326-2400  
;; TELEFAX: (415) 326-2422  
;;  
;; INFORMATION FOR SEQ ID NO: 74:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 132 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (oligonucleotide)  
;;  
US-08-474-040-74

Query Match 57.6%; Score 19; DB 1; Length 132;  
Best Local Similarity 81.5%; Fred. No. 19;  
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 TCTAGACCACCATGGGAGAGCGGAA 29  
|||||  
Db 5 TCTAGACCACCATGGGAGAGCGGAA 31

RESULT 13  
US-08-487-200-74  
; Sequence 74, Application US/08487200  
; Patent No. 5693762  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: CO. Man Sung  
; APPLICANT: SCHNEIDER, William P.  
; APPLICANT: LANDOLEI, Nicholas F.  
; APPLICANT: COELINGH, Kathleen L.  
; APPLICANT: SELICK, Harold E.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/487,200  
;; FILING DATE: 7-JUN-1995  
;; CLASSIFICATION: 424  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/634,278  
;; FILING DATE: 19-DEC-1990  
;; APPLICATION NUMBER: US 07/590,274  
;; FILING DATE: 28-SEP-1990  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/310,252  
;; FILING DATE: 13-FEB-1989  
;; APPLICATION NUMBER: US 07/290,975  
;; FILING DATE: 28-DEC-1988  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Smith, William M  
;; REGISTRATION NUMBER: 30,223  
;; REFERENCE/DOCKET NUMBER: 11823-002610  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 326-2400  
;; TELEFAX: (415) 326-2422  
;;  
;; INFORMATION FOR SEQ ID NO: 74:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 132 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (oligonucleotide)  
;;  
US-08-487-200-74

Query Match 57.6%; Score 19; DB 1; Length 132;  
Best Local Similarity 81.5%; Fred. No. 19;  
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 TCTAGACCACCATGGGAGAGCGGAA 29  
|||||  
Db 5 TCTAGACCACCATGGGAGAGCGGAA 31

RESULT 14  
US-08-484-537-74  
; Sequence 74, Application US/08484537  
; Patent No. 6180370  
; GENERAL INFORMATION:  
; APPLICANT: QUEEN, Cary L.  
; APPLICANT: CO. Man Sung  
; APPLICANT: SCHNEIDER, William P.  
; APPLICANT: LANDOLEI, Nicholas F.  
; APPLICANT: COELINGH, Kathleen L.  
; APPLICANT: SELICK, Harold E.  
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,537  
; FILING DATE:

```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 19-DEC-1990
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (oligonucleotide)
US-08-484-537-74

Query Match 57.6%; Score 19; DB 3; Length 132;
Best Local Similarity 81.5%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 TCTAGACCACCATGGGAGAGGCGGAA 29
| | | | | | | | | | | | | | | |
Db 5 TCTAGACCACCATGGGAGAGGCGGAA 31

RESULT 15
US-07-634-278-94
; Sequence 94, Application US/07634278
; Patent No. 5530101
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/634,278
; FILING DATE: 19-DEC-1990
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (oligonucleotide)
US-07-634-278-94

Query Match 57.6%; Score 19; DB 1; Length 134;
Best Local Similarity 81.5%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 TCTAGACCACCATGGGAGAGGCGGAA 29
| | | | | | | | | | | | | | | |
Db 5 TCTAGACCACCATGGGAGAGGCGGAA 31

Search completed: December 3, 2003, 14:51:44
Job time : 40.2727 secs
```



Publication No. US20030119104A1  
GENERAL INFORMATION: Perkins, Edward  
APPLICANT: Perkins, Edward  
APPLICANT: Perez, Carl  
APPLICANT: Lindenbaum, Michael  
APPLICANT: Greene, Amy  
APPLICANT: Leung, Josephine  
APPLICANT: Fleming, Elena  
APPLICANT: Stewart, Sandra  
APPLICANT: Shellard, Joan  
TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
FILE REFERENCE: 24601-420  
CURRENT APPLICATION NUMBER: US/10/161,403  
CURRENT FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: 60/294,758  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 60/366,891  
PRIOR FILING DATE: 2002-03-21  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 35  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Primer: Lamint 1  
US-10-161-403-3

Query Match 63.6%; Score 21; DB 14; Length 35;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 CATGGGAAGAGCGGAAGTCA 33  
DB 9 CATGGGAAGAGCGGAAGTCA 29

RESULT 3  
US-10-161-403-112  
Sequence 112, Application US/10161403  
Publication No. US20030119104A1  
GENERAL INFORMATION: Perkins, Edward  
APPLICANT: Perkins, Edward  
APPLICANT: Perez, Carl  
APPLICANT: Lindenbaum, Michael  
APPLICANT: Greene, Amy  
APPLICANT: Leung, Josephine  
APPLICANT: Fleming, Elena  
APPLICANT: Stewart, Sandra  
APPLICANT: Shellard, Joan  
TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
FILE REFERENCE: 24601-420  
CURRENT APPLICATION NUMBER: US/10/161,403  
CURRENT FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: 60/294,758  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 60/366,891  
PRIOR FILING DATE: 2002-03-21  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 112  
LENGTH: 5855  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: pCX-LamIntr Plasmid  
US-10-161-403-112

Query Match 63.6%; Score 21; DB 14; Length 5855;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 CATGGGAAGAGCGGAAGTCA 33

DB 1724 CATGGGAAGAGCGGAAGTCA 1744

RESULT 4  
US-10-161-403-127  
Sequence 127, Application US/10161403  
Publication No. US20030119104A1  
GENERAL INFORMATION: Perkins, Edward  
APPLICANT: Perkins, Edward  
APPLICANT: Perez, Carl  
APPLICANT: Lindenbaum, Michael  
APPLICANT: Greene, Amy  
APPLICANT: Leung, Josephine  
APPLICANT: Fleming, Elena  
APPLICANT: Stewart, Sandra  
APPLICANT: Shellard, Joan  
TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
FILE REFERENCE: 24601-420  
CURRENT APPLICATION NUMBER: US/10/161,403  
CURRENT FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: 60/294,758  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 60/366,891  
PRIOR FILING DATE: 2002-03-21  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 127  
LENGTH: 5855  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: pCX-LamInt Plasmid (Wildtype Integrase)  
US-10-161-403-127

Query Match 63.6%; Score 21; DB 14; Length 5855;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 CATGGGAAGAGCGGAAGTCA 33  
DB 1724 CATGGGAAGAGCGGAAGTCA 1744

RESULT 5  
US-10-027-632-129494/c  
Sequence 129494, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION: Wang, David G.  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome  
FILE REFERENCE: 108827,129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 129494  
LENGTH: 557

```

; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129494

Query Match      63.0%; Score 20.8; DB 12; Length 557;
Best Local Similarity 78.1%; Pred. No. 25;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1  GCTCTAGACCACCATGGGAGAGAGCGGAAGTC 32
          |||
Db      339  GCGCTGGGACAAACCTGGGAAAGCGAGGTGAAGTC 308
          |||

```

```

RESULT 6
US-10-027-632-129494/c
; Sequence 129494, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129494
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129494

```

```

RESULT 7
US-09-318-995-36667
; Sequence 36667, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/318,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36667
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

US-09-918-995-36667

Query Match      62.4%; Score 20.6; DB 11; Length 416;
Best Local Similarity 85.2%; Pred. No. 29;
Matches 23; Conservative 0; Mismatches 4; Indels 0

Qy      4 CTAGACGACCATGGGAAGCGGAAG 30
        |||||
Db      289 CTAGACGACCCCTGGGAAGAGGTGAAG 315

RESULT 8
US-09-918-995-12546
; Sequence 12546, Application US/09518995
; Publication No. US20030073623A1

```

```

RESULT 8
US-09-918-995-12546
; Sequence 12546, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12546
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-12546

```

```

RESULT 9
US-09-372-348-14
; Sequence 14, Application US/09372348A
; Publication No. US20030028005A1
; GENERAL INFORMATION:
; APPLICANT: Hedrick, Joseph A.
; APPLICANT: Wang, Luquan
; APPLICANT: Zlotnik, Albert
; APPLICANT: Murgolo, Nicholas J.
; APPLICANT: Greene, Jonathan R.
; APPLICANT: Johnston, James A.
; APPLICANT: Bazan, J. Fernando
; TITLE OF INVENTION: Mammalian Proteins; Related Reagents and Methods
; FILE REFERENCE: EX0884K(2d)
; CURRENT APPLICATION NUMBER: US/09/372,348A
; CURRENT FILING DATE: 1999-08-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 2244
; TYPE: DNA
; ORGANISM: primate
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2244)
US-09-372-348-14

```

```

Query Match      62.4%; Score 20.6; DB 11; Length 2244;
Best Local Similarity 85.2%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Ov      4 CTAGACCACCATGGGAAGAGGGCCAG 30

```



```

; Sequence 903, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 903
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-903

Query Match          60.6%; Score 20; DB 10; Length 285;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      14 ATGGGAAGAGCGCAAGTCA 33
        |||||||
Db       53 ATGGGAAGAGCGCAAGTCA 82

RESULT 13
US-09-867-701-917/c
; Sequence 917, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 917
; LENGTH: 331
; TYPE: DNA
; ORGANISM: Homo sapien
; NAME/KEY: misc_feature
; LOCATION: (1)...(331)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-917

Query Match          60.6%; Score 20; DB 10; Length 331;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      14 ATGGGAAGAGCGCAAGTCA 33
        |||||||
Db       242 ATGGGAAGAGCGCAAGTCA 223

RESULT 14
US-09-867-701-580/c
; Sequence 580, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497

```

; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 580  
; LENGTH: 340  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-867-701-580

Query Match 60.6%; Score 20; DB 10; Length 340;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 ATGGGAGAGCGCGAAGTCA 33  
Db 249 ATGGGAGAGCGCGAAGTCA 230

RESULT 15  
US-10-161-403-37  
; Sequence 37, Application US/10161403  
; Publication No. US20030119104A1  
; GENERAL INFORMATION:  
; APPLICANT: Perkins, Edward  
; APPLICANT: Perez, Carl  
; APPLICANT: Lindenbaum, Michael  
; APPLICANT: Greene, Amy  
; APPLICANT: Leung, Josephine  
; APPLICANT: Fleming, Elena  
; APPLICANT: Stewart, Sandra  
; APPLICANT: Shellard, Joan  
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
; FILE REFERENCE: 24601-420  
; CURRENT APPLICATION NUMBER: US/10/161,403  
; CURRENT FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: 60/294,758  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/366,891  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37  
; LENGTH: 1071  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Integrase E174R  
; NAME/KEY: CDS  
; LOCATION: (1)...(1071)  
; OTHER INFORMATION: Nucleotide sequence encoding Integrase E147R  
US-10-161-403-37

Query Match 60.6%; Score 20; DB 14; Length 1071;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 ATGGGAGAGCGCGAAGTCA 33  
Db 1 ATGGGAGAGCGCGAAGTCA 20

Search completed: December 3, 2003, 17:22:22  
Job time : 148.273 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 1479.82 Seconds  
(without alignments)  
541.991 Million cell updates/sec

Title: US-10-082-772-1  
Perfect score: 33  
Sequence: 1 gcttagaccaccatgggaagcggaagtcga 33

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hcc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21.8	66.1	462	14	N68492
2	21.2	64.2	944	10	BF540269
3	21.2	64.2	957	10	BF301050
4	21	63.6	401	9	AA622176

C	5	21	63.6	561	14	CB874845
C	6	20.8	63.0	442	9	AU182375
C	7	20.8	63.0	460	9	AU183692
C	8	20.8	63.0	491	9	AL138085
C	9	20.8	63.0	569	9	AU093457
C	10	20.8	63.0	581	14	C98571
C	11	20.8	63.0	585	10	BF047732
C	12	20.8	63.0	604	9	AU093866
C	13	20.8	63.0	657	9	AU094204
C	14	20.8	63.0	658	9	AU165497
C	15	20.8	63.0	709	9	AU165529
C	16	20.6	62.4	318	13	BQ326535
C	17	20.6	62.4	324	10	BF802501
C	18	20.6	62.4	337	12	BG982644
C	19	20.6	62.4	408	12	BM721654
C	20	20.6	62.4	420	12	BQ027270
C	21	20.6	62.4	477	9	AW503928
C	22	20.6	62.4	519	10	AW937877
C	23	20.6	62.4	616	10	BE153138
C	24	20.6	62.4	772	12	BG762801
C	25	20.6	62.4	864	10	BE797208
C	26	20.6	62.4	911	13	BH949030
C	27	20.6	62.4	940	13	BQ651615
C	28	20.6	62.4	956	13	BQ652033
C	29	20.6	62.4	980	29	CHS06X5E
C	30	20.6	62.4	3116	11	BC027992
C	31	20.6	62.4	3191	11	BC016487
C	32	20.4	61.8	351	9	AI779832
C	33	20.4	61.8	414	9	AW462957
C	34	20.4	61.8	443	9	AI483405
C	35	20.4	61.8	451	14	R14443
C	36	20.4	61.8	466	9	AW462380
C	37	20.4	61.8	498	10	BE750832
C	38	20.4	61.8	512	9	AW091564
C	39	20.4	61.8	520	12	B1206604
C	40	20.4	61.8	521	12	B1207570
C	41	20.4	61.8	563	9	AI776731
C	42	20.4	61.8	569	10	BG130217
C	43	20.4	61.8	647	12	BG826410
C	44	20.4	61.8	763	10	BF134304
C	45	20.4	61.8	1641	10	BF133376

#### ALIGNMENTS

RESULT 1  
LOCUS N68492 462 bp mRNA linear EST 13-MAR-1996  
DEFINITION zai4h04.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
IMAGE:292567.3', mRNA sequence.  
ACCESSION N68492  
VERSION N68492.1 GI:1224653  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 462)  
AUTHORS Hillier, L., Clark, N., Dubuque, T., Eliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.  
TITLE The WashU-Merck EST Project  
JOURNAL Unpublished  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.

[illegible]

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 401)  
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 348.

FEATURES  
source  
1. 401  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:1147830"  
/tissue\_type="colon tumor RER+"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP C09"  
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from RER+ colon tumor, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Ronaldo (Soares4)." . "

BASE COUNT 74 a 103 c 120 g 104 t  
ORIGIN

Query Match 63.6%; Score 21; DB 9; Length 401;  
Best Local Similarity 82.8%; Pred. No. 8.5e+02;  
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCTCTAGACCACTGGGAGAGCGGAA 29  
|||||  
Db 80 GCTCTGGTCCACCATGGCGCGGCGGA 108  
|||||

RESULT 5  
CB874845/c  
LOCUS CB874845 561 bp mRNA linear EST 23-APR-2003  
DEFINITION HK06H21w HK Hordeum vulgare subsp. vulgare cDNA clone HK06H21  
3-PRIME, mRNA sequence.  
ACCESSION CB874845  
VERSION CB874845.1 GI:30076831  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Triticeae; Hordeum.  
REFERENCE 1 (bases 1 to 561)  
AUTHORS Zhang, H., Weschke, W., Michalek, W., Stein, N. and Graner, A.  
TITLE EST sequencing and analysis in barley (2002)  
JOURNAL Unpublished  
COMMENT Contact: Stein Nils  
Molecular Markers Group, Department Genbank  
Institute of Plant Genetics and Crop Plant Research (IPK)  
Corrensstr. 3, 06466, Gatersleben, Germany

Tel: 039482-5522  
Fax: 039482-5595  
Email: stein@ipk-gatersleben.de  
Insert Length: 561 Std Error: 0.00  
Place: 6 row: H column: 21  
Seq primer: 77.  
Location/Qualifiers  
1. 561  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="barke"  
/db\_xref="GABI:561191"  
/db\_xref="taxon:112509"  
/clone="HX06H21"  
/tissue\_type="apex"  
/dev\_stage="apex (3-5 mm in size)"  
/lab\_host="XL10-Gold"  
/clone\_lib="HX"  
/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."  
BASE COUNT 148 a 127 c 112 g 174 t  
ORIGIN

Query Match 63.6%; Score 21; DB 14; Length 561;  
Best Local Similarity 82.8%; Pred. No. 9.5e+02;  
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TAGACCACTGGGAGAGCGGAGTCA 33  
|||||  
Db 397 TAAACCATCTTGGGAGAGCGGAGTCA 369  
|||||

RESULT 6  
AU182375/c  
LOCUS AU182375 442 bp mRNA linear EST 03-APR-2002  
DEFINITION AU182375 Rice panicle at flowering stage Oryza sativa (japonica cultivar-group) cDNA clone E2576, mRNA sequence.  
ACCESSION AU182375  
VERSION AU182375.1 GI:13896039  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 442)  
AUTHORS Sasaki, T. and Yamamoto, K.  
TITLE Rice cDNA from panicle at flowering stage (2001)  
JOURNAL Unpublished  
COMMENT Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
PROJECT = 'RGP'.  
Location/Qualifiers  
1. 442  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="E2576"  
/dev\_stage="flowering stage"  
/clone\_lib="Rice panicle at flowering stage"  
/note="Organ: panicle; Rice cDNA from panicle at flowering

```

REFERENCE 1 (bases 1 to 491)
AUTHORS Ansgore, W., Wirkner, U., Mewes, W., Weil, B. and Wiemann, S.
TITLE EST (Ansgore, W., Wirkner, U., Mewes, H.W., Weil, B. and Wiemann, S.)
JOURNAL Unpublished
COMMENT Contact: Ansgore W
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the CDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp761P2315) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             source
    Location/Qualifiers
        1..491
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp761P2315"
            /tissue_type="amygdala"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="761 (synonym: hamy2)"
            /note="vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      87 a 136 c 127 g 141 t
ORIGIN
Query Match      63.0%; Score 20.8; DB 9; Length 491;
Best Local Similarity 78.1%; Pred. No. 1.le+03;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GCTTAGACCACTGGAGAGAGGCGGAGTC 32
|||||
DB 471 GCTGAGACCACTGGAGAGAGGCGGAAATC 440
|||||

RESULT 9
AU093457/c
LOCUS AU093457.1 linear EST 03-APR-2002
DEFINITION AU093457 Rice panicle at flowering stage Oryza sativa (japonica
cultivar-group) cDNA clone E1031, mRNA sequence.
ACCESSION AU093457
VERSION AU093457.1 GI:8789087
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 569)
Sasaki, T. and Yamamoto, K.
Rice cDNA from panicle at flowering stage (2000)
Unpublished
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT = RGP.
FEATURES             source
    Location/Qualifiers
        1..569
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /cultivar="Nipponbare"
            /db_xref="taxon:39947"
            /clone="E1031"
            /dev_stage="flowering stage"

```

```

/clone_lib="Rice panicle at flowering stage"
/note="Organ: panicle; Rice cDNA from panicle at flowering
stage"
BASE COUNT      147 a   124 c   106 g   132 t
ORIGIN

Query Match      63.0%; Score 20.8; DB 9; Length 569;
Best Local Similarity 78.1%; Pred. No. 1.1e+03;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CTCCTAGACACCATGGGAAGGCGAAGTCA 33
Db 200 CGCTGGACCATCGTGGTAGAAGGGGAATGA 169

RESULT 10
C98571/c
LOCUS
DEFINITION      C98571 Rice panicle at flowering stage Oryza sativa (japonica
cultiivar-group) cDNA clone E0419_6Z, mRNA sequence.
ACCESSION      C98571
VERSION        C98571.1 GI:3761323
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
AUTHORS        1 (bases 1 to 581)
TITLE          Rice cDNA from panicle at flowering stage
JOURNAL
COMMENT        Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT = 'RGP'
POLYA=No.

FEATURES
source
Location/Qualifiers
1..581
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39947"
/clone="E0419_6Z"
/dev_stage="flowering stage"
/clone_lib="Rice panicle at flowering stage"
/note="Organ: panicle; Rice cDNA from panicle at flowering
stage"
BASE COUNT      145 a   139 c   112 g   184 t
ORIGIN

Query Match      63.0%; Score 20.8; DB 14; Length 581;
Best Local Similarity 78.1%; Pred. No. 1.1e+03;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CTCCTAGACACCATGGGAAGGCGAAGTCA 33
Db 271 CGCTGGACCATCGTGGTAGAAGGGGAATGA 240

RESULT 11
BF047732
LOCUS
DEFINITION      BF047732 585 bp mRNA EST 11-OCT-2000
dc82a02.y1 NICHDXGCG 001 Xenopus laevis cDNA clone IMAGE3:3403562 5',
similar to SW-SPCO_MOUSE Q62251 SPECTRIN BETA CHAIN, BRAIN ;, mRNA
sequence.
ACCESSION      BF047732
VERSION        BF047732.1 GI:10766235
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodidae; Xenopus.
Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodidae; Xenopus.
1 (bases 1 to 585)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Other ESTs: dc82a02.x1
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 384.
Location/Qualifiers
1..585
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:3403562"
/tissue_type="oocytes"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDXGCG 001"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.2 kb. Constructed by Life Technologies."
BASE COUNT      158 a   151 c   151 g   125 t
ORIGIN

Query Match      53.0%; Score 20.8; DB 10; Length 585;
Best Local Similarity 78.1%; Pred. No. 1.1e+03;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CTCCTAGACACCATGGGAAGGCGAAGTCA 33
Db 412 CTCAGCCCATCGTGGGAGATGCGTATCCA 443

RESULT 12
AU093866/c
LOCUS
DEFINITION      AU093866 Rice panicle at flowering stage Oryza sativa (japonica
cultiivar-group) cDNA clone E0981, mRNA sequence.
ACCESSION      AU093866
VERSION        AU093866.1 GI:8856548
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 604)
Sasaki, T. and Yamamoto, K.
Rice cDNA from panicle at flowering stage (2000)
Unpublished
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT = 'RGP'
Location/Qualifiers
1..604
/organism="Oryza sativa (japonica cultivar-group)"

```

```

/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E0981"
/dev_stage="flowering stage"
/clone_lib="Rice panicle at flowering stage"
/notes="Organ: panicle; Rice cDNA from panicle at flowering stage"
BASE COUNT      150 a      147 c      116 g      190 t      1 others
ORIGIN

Query Match      63.0%; Score 20.8; DB 9; Length 604;
Best Local Similarity 78.1%; Pred. No. 1.2e+03;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      2      CTCGTAGACCACCATGGGAAGAGCGGAAGTCA 33
DB      297      CGCTGGACCACTTGTGTAGAGGGGAATGA 266

RESULT 13
AU094204/c
LOCUS
DEFINITION
AU094204 Rice panicle at flowering stage Oryza sativa (japonica cultivar-group) cDNA clone E3633, mRNA sequence.
AU094204
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 657)
AUTHORS
Sasaki,T. and Yamanoto,K.
TITLE
Rice cDNA from panicle at flowering stage (2000)
JOURNAL
Unpublished
COMMENT
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602 Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/PROJECT="RGP".
FEATURES
Location/Qualifiers
1..657
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E3633"
/dev_stage="flowering stage"
/clone_lib="Rice panicle at flowering stage"
/notes="Organ: panicle; Rice cDNA from panicle at flowering stage"
BASE COUNT      165 a      192 c      142 g      157 t      1 others
ORIGIN

Query Match      63.0%; Score 20.8; DB 9; Length 657;
Best Local Similarity 78.1%; Pred. No. 1.2e+03;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      2      CTCGTAGACCACCATGGGAAGAGCGGAAGTCA 33
DB      497      CGCTGGACCACTTGTGTAGAGGGGAATGA 466

RESULT 14
AU165497/c
LOCUS
DEFINITION
AU165497 Rice panicle at flowering stage Oryza sativa (japonica cultivar-group) cDNA clone E1731, mRNA sequence.

```

ACCESSION AU165497  
 VERSION AU165497.1 GI:11564861  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1 (bases 1 to 658)  
 REFERENCE Sasaki, T. and Yamamoto, K.  
 AUTHORS Rice cDNA from panicle at flowering stage (2000)  
 TITLE Unpublished  
 JOURNAL  
 COMMENT Contact: Takuji Sasaki  
 National Institute of Agrobiological Resources  
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
 305-8602, Japan  
 Tel: 81-298-38-7441  
 Fax: 81-298-38-7468  
 Email: tsasaki@ab.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
 PROJECT = 'RGP',

FEATURES  
 source  
 1..658  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="E1731"  
 /dev\_stage="flowering stage"  
 /clone\_11b="Rice panicle at flowering stage"  
 /note="Organ: panicle; Rice cDNA from panicle at flowering stage"  
 BASE COUNT 156 a 171 c 131 g 199 t 1 others  
 ORIGIN

Query Match 63.0%; Score 20.8; DB 9; Length 658;  
 Best Local Similarity 78.1%; Pred. No. 1.2e+03;  
 Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CTCGTAGACCACCATGGGAAGAGCGGAAGTCA 33  
 Db 374 CGCTGACCAACCATGGTAGAAGGGAATGA 343

RESULT 15  
 AU165529/c  
 LOCUS AU165529 709 bp mRNA linear EST 03-APR-2002  
 DEFINITION AU165529 Rice panicle at flowering stage Oryza sativa (japonica cultivar-group) cDNA clone E2154, mRNA sequence.  
 ACCESSION AU165529  
 VERSION AU165529.1 GI:11564893  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1 (bases 1 to 709)  
 REFERENCE Sasaki, T. and Yamamoto, K.  
 AUTHORS Rice cDNA from panicle at flowering stage (2000)  
 TITLE Unpublished  
 JOURNAL  
 COMMENT Contact: Takuji Sasaki  
 National Institute of Agrobiological Resources  
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
 305-8602, Japan  
 Tel: 81-298-38-7441  
 Fax: 81-298-38-7468  
 Email: tsasaki@ab.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
 PROJECT = 'RGP',

FEATURES  
 source  
 1..709  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"



Search completed: December 3, 2003, 14:49:13  
Job time : 1483.82 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 1087.6 Seconds  
(without alignments)  
1504.579 Million cell updates/sec

Title: US-10-082-772-2

Perfect score: 40

Sequence: 1 aaggaaagcgccgcattattgatttcatttggcc 40

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sv.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	62.5	33	6	AR160204
2	25	62.5	33	6	E36467
3	25	62.5	201	6	AR044609
4	25	62.5	201	6	E05439
5	25	62.5	201	6	I113139
6	25	62.5	201	6	I36498
7	25	62.5	260	7	SIPOLYL2
8	25	62.5	361	7	LAMINTATT
9	25	62.5	368	7	LAVATTBOP
10	25	62.5	380	8	AB026619
11	25	62.5	717	9	HS332250
12	25	62.5	979	9	HUMDTN02
13	25	62.5	1090	9	HS3325746
14	25	62.5	1421	10	AF121976
15	25	62.5	1423	3	PPGFRA
16	25	62.5	1668	9	MACHSS
17	25	62.5	2408	6	AR059581
18	25	62.5	2408	6	AR063063
19	25	62.5	2408	6	I93548
20	25	62.5	2616	7	P434XISINT
21	25	62.5	2758	3	PFARPC
22	25	62.5	2934	3	TRBGP85B
23	25	62.5	2959	12	AY048723
24	25	62.5	3049	6	AX430197
25	25	62.5	3049	6	AX449167
26	25	62.5	3111	8	CRPERAS
27	25	62.5	3423	5	CHKPLB2
28	25	62.5	3484	6	A91914
29	25	62.5	3484	6	AR309120
30	25	62.5	3484	6	BD007602
31	25	62.5	3485	12	AF178449
32	25	62.5	3485	12	AF178450
33	25	62.5	3663	12	AY048716
34	25	62.5	3695	12	AY048733
35	25	62.5	3757	6	AR080463
36	25	62.5	3757	10	MUSPROT6A
37	25	62.5	4105	12	AF271663
38	25	62.5	4190	12	XXU13848
39	25	62.5	4549	12	AF178452
40	25	62.5	4549	12	AF178453
41	25	62.5	4782	12	AY048740
42	25	62.5	5129	12	AY048741
43	25	62.5	5201	12	AY048715
44	25	62.5	5349	6	A71437
45	25	62.5	5349	6	AR207455

## ALIGNMENTS

RESULT 1	AR160204	Sequence 5	from patent US 6255048.	33 bp	DNA	linear	PAT 17-OCT-2001
LOCUS	AR160204	Sequence 5	from patent US 6255048.	33 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	AR160204	Sequence 5	from patent US 6255048.	33 bp	DNA	linear	PAT 17-OCT-2001
ACCESSION	AR160204	Sequence 5	from patent US 6255048.	33 bp	DNA	linear	PAT 17-OCT-2001
VERSION	AR160204.1	GI:16223838					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 33)						
AUTHORS	Hosoi,S., Kojima,M. and Kadouchi,S.						
TITLE	Highly sensitive Fluoroassay						
JOURNAL	Patent: US 6255048-A 5 03-JUL-2001;						
FEATURES	Location/Qualifiers						

```
source
1. .33
/organism="unknown"
BASE COUNT      7 a      9 c      2 g      15 t
ORIGIN

Query Match      62.5%; Score 25; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTCGAATTTGTCC 40
Db 2 TCATTATTGATTCGAATTTGTCC 26

RESULT 2
E36467
LOCUS      E36467      33 bp      DNA      linear      PAT 18-JUN-2001
DEFINITION DNA polymerase having improved heat stability and elevated length
and efficiency of primer extension.
ACCESSION   E36467
VERSION     E36467.1 GI:13022679
KEYWORDS    JP 1999239492-A/22.
SOURCE      Bacteriophage lambda
ORGANISM    Bacteriophage lambda
Virus; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
Lambda-like viruses.
1 (bases 1 to 33)
REFERENCE   Barnes,U.M.
AUTHORS     Barnes,U.M.
TITLE       DNA polymerase having improved heat stability and elevated length
and efficiency of primer extension
JOURNAL     Patent: JP 1999239492-A 22 07-SEP-1999;
TAKARA SHUZO CO LTD
COMMENT     OS bacteriophage lambda
FN JP 1999239492-A/22
PD 07-SEP-1999
PF 17-DEC-1998 JP 1998359199
PR 19-FEB-1993 US 08/021623
PI BARNES UEIN M
PC C12N15/09,C12N9/12,C12Q1/68,C12N15/00
CC
FH
FT
FEATURES
source
1. .33
Location/Qualifiers
/organism="Bacteriophage lambda"
/mol_type="genomic DNA"
/db_xref="taxon:10710"
BASE COUNT      7 a      9 c      2 g      15 t
ORIGIN

Query Match      62.5%; Score 25; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTCGAATTTGTCC 40
Db 2 TCATTATTGATTCGAATTTGTCC 26

RESULT 3
AR044609
LOCUS      AR044609      201 bp      DNA      linear      PAT 29-SEP-1999
DEFINITION Sequence 18 from patent US 5817506.
ACCESSION   AR044609
VERSION     AR044609.1 GI:5966074
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 201)
AUTHORS     Okano,K. and Kambara,H.
TITLE       Polynucleotide capturing support for capturing, eluting and
collecting polynucleotides in a sample solution
Patent: US 5817506-A 18 06-OCT-1998;
Location/Qualifiers
1. .201
/organism="unknown"
BASE COUNT      58 a      38 c      27 g      78 t
ORIGIN

Query Match      62.5%; Score 25; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTCGAATTTGTCC 40
Db 130 TCATTATTGATTCGAATTTGTCC 154

RESULT 4
E05439
LOCUS      E05439      201 bp      DNA      linear      PAT 29-SEP-1997
DEFINITION Oligonucleotide.
ACCESSION   E05439
VERSION     E05439.1 GI:2173628
KEYWORDS    JP 1993236997-A/11.
SOURCE      synthetic construct
ORGANISM    artificial sequences.
1 (bases 1 to 201)
REFERENCE   Okano,K. and Kambara,H.
AUTHORS     Okano,K. and Kambara,H.
TITLE       CHIP FOR CATCHING POLYNUCLEOTIDE
JOURNAL     Patent: JP 1993236997-A 11 17-SEP-1993;
HITACHI LTD
COMMENT     OS Artificial gene
CC Artificial sequence; Genes.
FN JP 1993236997-A/11
PD 17-SEP-1993
PF 28-FEB-1992 JP 1992042829
PI OKANO KAZUNOBU, KANBARA HIDEKI
PC C12Q1/68;
CC strandedness: Single;
topology: linear.
FEATURES
source
1. .201
Location/Qualifiers
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT      58 a      38 c      27 g      78 t
ORIGIN

Query Match      62.5%; Score 25; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTCGAATTTGTCC 40
Db 130 TCATTATTGATTCGAATTTGTCC 154

RESULT 5
I13139
LOCUS      I13139      201 bp      DNA      linear      PAT 26-JUL-1995
DEFINITION Sequence 18 from patent US 5434049.
ACCESSION   I13139
VERSION     I13139.1 GI:910488
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 201)
AUTHORS     Okano,K. and Kambara,H.
TITLE       Separation of polynucleotides using supports having a plurality of
electrode-containing cells
Patent: US 5434049-A 18 18-JUL-1995;
JOURNAL
```

```
JOURNAL
collecting polynucleotides in a sample solution
Patent: US 5817506-A 18 06-OCT-1998;
Location/Qualifiers
1. .201
/organism="unknown"
BASE COUNT      58 a      38 c      27 g      78 t
ORIGIN

Query Match      62.5%; Score 25; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTCGAATTTGTCC 40
Db 130 TCATTATTGATTCGAATTTGTCC 154

RESULT 4
E05439
LOCUS      E05439      201 bp      DNA      linear      PAT 29-SEP-1997
DEFINITION Oligonucleotide.
ACCESSION   E05439
VERSION     E05439.1 GI:2173628
KEYWORDS    JP 1993236997-A/11.
SOURCE      synthetic construct
ORGANISM    artificial sequences.
1 (bases 1 to 201)
REFERENCE   Okano,K. and Kambara,H.
AUTHORS     Okano,K. and Kambara,H.
TITLE       CHIP FOR CATCHING POLYNUCLEOTIDE
JOURNAL     Patent: JP 1993236997-A 11 17-SEP-1993;
HITACHI LTD
COMMENT     OS Artificial gene
CC Artificial sequence; Genes.
FN JP 1993236997-A/11
PD 17-SEP-1993
PF 28-FEB-1992 JP 1992042829
PI OKANO KAZUNOBU, KANBARA HIDEKI
PC C12Q1/68;
CC strandedness: Single;
topology: linear.
FEATURES
source
1. .201
Location/Qualifiers
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT      58 a      38 c      27 g      78 t
ORIGIN

Query Match      62.5%; Score 25; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTCGAATTTGTCC 40
Db 130 TCATTATTGATTCGAATTTGTCC 154

RESULT 5
I13139
LOCUS      I13139      201 bp      DNA      linear      PAT 26-JUL-1995
DEFINITION Sequence 18 from patent US 5434049.
ACCESSION   I13139
VERSION     I13139.1 GI:910488
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 201)
AUTHORS     Okano,K. and Kambara,H.
TITLE       Separation of polynucleotides using supports having a plurality of
electrode-containing cells
Patent: US 5434049-A 18 18-JUL-1995;
JOURNAL
```

FEATURES  
source  
Location/Qualifiers  
1. .201  
/organism="unknown"  
BASE COUNT 58 a 38 c 27 g 78 t  
ORIGIN

Query Match 62.5%; Score 25; DB 6; Length 201;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTGTGCC 40  
|||||  
Db 130 TCATTATTGATTTCAATTGTGCC 154  
|||||

RESULT 6  
I36498  
LOCUS I36498 201 bp DNA linear PAT 13-MAY-1997  
DEFINITION Sequence 18 from patent US 5607646.  
ACCESSION I36498  
VERSION I36498.1 GI:2086323  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE  
1 (bases 1 to 201)  
Okano,K. and Kanbara,H.  
TITLE Device for separating polynucleotides having a plurality of  
electrode-containing cells and movable collecting capillary  
JOURNAL Patent: US 5607646-A 18 04-MAR-1997;  
FEATURES Location/Qualifiers  
source 1. .201  
/organism="unknown"  
BASE COUNT 58 a 38 c 27 g 78 t  
ORIGIN

Query Match 62.5%; Score 25; DB 6; Length 201;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTGTGCC 40  
|||||  
Db 130 TCATTATTGATTTCAATTGTGCC 154  
|||||

RESULT 7  
SIPOLYL2  
LOCUS SIPOLYL2 260 bp DNA linear PHG 14-FEB-1995  
DEFINITION Bacteriophage lambda EMBL3 left arm DNA with int gene, partial cds.  
ACCESSION X58668  
VERSION X58668.1 GI:15757  
KEYWORDS int gene; integrase; polylinker; vector.  
SOURCE Bacteriophage lambda  
ORGANISM Bacteriophage lambda  
Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
Lambda-like viruses.

REFERENCE  
1  
Manninen,I. and Schulman,A.H.  
AUTHORS  
TITLE The lambda EMBL3 polylinker and surrounding region for PCR primers  
JOURNAL Biotechniques 14 (2), 174 (1993)  
MEDLINE 93159764  
PUBMED 8381652

REFERENCE  
2 (bases 1 to 260)  
Schulman,A.H.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (27-MAR-1991) A.H. Schulman, Institute of Biotechnology,  
University of Helsinki P.O. Box 45 (Karvaamokuja 3a), FIN-00014  
Helsinki, FINLAND  
COMMENT See M17233, N.A.R. 8:1765-1782(1980) and N.A.R. 16:1633(1988) for  
overlapping sequences.  
See X58667 for related sequence.  
FEATURES Location/Qualifiers  
source 1. .260

/organism="Bacteriophage lambda"  
/mol\_type="genomic DNA"  
/strain="EMBL3"  
/db\_xref="taxon:10710"  
complement(<1..237)  
partial"  
/note="int gene  
misc\_feature  
complement(49..74)  
/note="int binding site"  
misc\_feature 238..260  
/note="polylinker"  
/evidence=experimental  
BASE COUNT 71 a 59 c 54 g 76 t  
ORIGIN

Query Match 62.5%; Score 25; DB 7; Length 260;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTGTGCC 40  
|||||  
Db 70 TCATTATTGATTTCAATTGTGCC 94  
|||||

RESULT 8  
LAMINTATT/C  
LOCUS LAMINTATT 361 bp DNA linear PHG 28-APR-1993  
DEFINITION Bacteriophage lambda int gene region.  
ACCESSION M23841  
VERSION M23841.1 GI:215177  
KEYWORDS site-specific recombination.  
SOURCE Bacteriophage lambda  
ORGANISM Bacteriophage lambda  
Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
Lambda-like viruses.

REFERENCE  
1 (bases 1 to 361)  
Schmeissner,U., McKenney,K., Rosenberg,M. and Court,D.  
AUTHORS Removal of a terminator structure by RNA processing regulates int  
gene expression  
JOURNAL J. Mol. Biol. 176 (1), 39-53 (1984)  
MEDLINE 84242838  
PUBMED 6234400

COMMENT Original source text: Bacteriophage lambda DNA.  
FEATURES Location/Qualifiers  
source 1. .361  
/organism="Bacteriophage lambda"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10710"  
BASE COUNT 134 a 60 c 60 g 107 t  
ORIGIN

Query Match 62.5%; Score 25; DB 7; Length 361;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTGTGCC 40  
|||||  
Db 31 TCATTATTGATTTCAATTGTGCC 7  
|||||

RESULT 9  
LAMATTBOP  
LOCUS LAMATTBOP 368 bp DNA linear PHG 28-APR-1993  
DEFINITION BOP' sequence from lambda-rifd18, including E.coli insert with  
secondary att site near E.coli rrnB operon.  
ACCESSION L00023 J01640  
VERSION L00023.1 GI:215101  
KEYWORDS ATT site; ATT-lambda site; attachment site; rrnB operon.  
SOURCE Bacteriophage lambda  
ORGANISM Bacteriophage lambda  
Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
Lambda-like viruses.

REFERENCE  
1 (bases 1 to 368)

AUTHORS Csordas-Toth, E., Boros, I. and Venetianer, P.  
 TITLE Nucleotide sequence of a secondary attachment site for  
 JOURNAL Bacteriophage lambda on the Escherichia coli chromosome  
 MEDLINE Nucleic Acids Res. 7 (5), 1335-1341 (1979)  
 PUBMED 80079294  
 COMMENT 160033  
 Original source text: Lambda-rif<sup>r</sup> transducing phage with E.coli  
 insertion.  
 The attachment site sequences on E.coli (attB) and lambda (attP)  
 contain a 15 bp long complete homology termed common core. Compared  
 with Grantham's data.

FEATURES  
 Source  
 1. 368  
 /organism="Bacteriophage lambda"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10710"  
 115. .129  
 /note="common core region secondary att site"  
 complement(203. .368)  
 /partial  
 /note="integration protein (int)"  
 /codon\_start=2  
 /transl\_table=11  
 /protein\_id="AAA32242.1"  
 /db\_xref="GI:215102"  
 /translation="DPTPHLRSARLYEKQISDKFAQLHGHKSDTMSQYRDDR  
 GREWDKIEIK"  
 BASE COUNT 99 a 81 c 80 g 108 t  
 ORIGIN  
 40 bp upstream from AluI site.

Query Match 62.5%; Score 25; DB 7; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCGAATTTGTC 40  
 |||||  
 Db 200 TCATTATTGATTCGAATTTGTC 224  
 |||||

RESULT 10  
 LOCUS AB026619 380 bp DNA linear PLN 22-OCT-2002  
 DEFINITION Nicotiana sylvestris Nsp3 gene for phosphoenolpyruvate  
 carboxylase, partial cds.  
 ACCESSION AB026619  
 VERSION AB026619.1 GI:20152184  
 KEYWORDS phosphoenolpyruvate carboxylase.  
 SOURCE Nicotiana sylvestris (wood tobacco)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE  
 1 Koizumi, N., Toyota, K., Kitajima, S., Yamada, Y. and Sato, F.  
 AUTHORS  
 TITLE Genomic structure and promoter analysis of phosphoenolpyruvate  
 carboxylase in a C3 plant, Nicotiana sylvestris  
 JOURNAL Biosci. Biotechnol. Biochem. 66 (8), 1691-1696 (2002)  
 MEDLINE 2240621  
 PUBMED 12353629

REFERENCE  
 2 (bases 1 to 380)  
 AUTHORS Sato, F. and Koizumi, N.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-APR-1999) Fumihiko Sato, Kyoto University, Department  
 of Plant Gene and Totipotency, Kitashirakawa, Sakyo, Kyoto, Kyoto  
 606-8502, Japan (E-mail:fumihiko@kais.kyoto-u.ac.jp,  
 URL:http://callus.kais.kyoto-u.ac.jp/callus.html,  
 Tel:81-75-753-6381, Fax:81-75-753-6398)

FEATURES  
 Location/Qualifiers  
 1. 380  
 /organism="Nicotiana sylvestris"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4096"  
 /clone\_lib="EMBL3"

gene 321. .380  
 /gene="Nsp3c3"  
 CDS 321. .>380  
 /gene="Nsp3c3"  
 /codon\_start=1  
 /product="phosphoenolpyruvate carboxylase"  
 /protein\_id="BAB9365.1"  
 /db\_xref="GI:20152185"  
 /translation="MATRNLEKLASIDAQLRALV"  
 BASE COUNT 105 a 75 c 85 g 115 t  
 ORIGIN

Query Match 62.5%; Score 25; DB 8; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCGAATTTGTC 40  
 |||||  
 Db 58 TCATTATTGATTCGAATTTGTC 82  
 |||||

RESULT 11  
 LOCUS HSA332250/c 717 bp DNA linear PRI 18-JUL-2002  
 DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone  
 N86-767S.  
 ACCESSION AJ332250  
 VERSION AJ332250.1 GI:15876668  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 (bases 1 to 717)  
 AUTHORS Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F., Kvasha, S.M.,  
 Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V.,  
 Levitsky, V.G., Kolchanov, N.A., Protodopov, A.I., Kashuba, V.I.,  
 Kisselev, L.L., Wasserman, W., Wahlestedt, C. and Zabarovsky, E.R.  
 TITLE NotI flanking sequences: a tool for gene discovery and verification  
 of the human genome  
 JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)  
 MEDLINE 22131767  
 PUBMED 12136098

REFERENCE  
 2 (bases 1 to 717)  
 AUTHORS Zabarovsky, E.R.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,  
 Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,  
 Sweden

FEATURES  
 Location/Qualifiers  
 1. 717  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="N86-767S"  
 BASE COUNT 194 a 192 c 164 g 167 t  
 ORIGIN

Query Match 62.5%; Score 25; DB 9; Length 717;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCGAATTTGTC 40  
 |||||  
 Db 133 TCATTATTGATTCGAATTTGTC 109  
 |||||

RESULT 12  
 LOCUS HUMDTN02 979 bp DNA linear PRI 21-NOV-2001  
 DEFINITION Human dystrobrein (DTN) gene, exon 2.  
 ACCESSION U84530  
 VERSION U84530.1 GI:2149295

KEYWORDS  
SEGMENT 2 of 23  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 979)

AUTHORS Sadoulet-Puccio,H.M., Feener,C.A., Schaid,D.J., Thibodeau,S.N., Michels,V.V. and Kunkel,L.M.

TITLE The genomic organization of human dystrobrevin

JOURNAL Neurogenetics 1 (1), 37-42 (1997)

MEDLINE 20197320

PUBMED 10735273

REFERENCE 2 (bases 1 to 979)

AUTHORS Sadoulet-Puccio,H.M., Feener,C.A. and Kunkel,L.M.

TITLE Direct Submission

JOURNAL Submitted (09-JAN-1997) Genetics, HHMI, 320 Longwood Avenue, Boston, MA 02115, USA

FEATURES  
source  
1. .979  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18q12"  
450..530  
/gene="DTN"  
/number=2  
a 199 c 183 g 296 t

exon

BASE COUNT 301 a 199 c 183 g 296 t

ORIGIN

Query Match 62.5%; Score 25; DB 9; Length 979;  
Best Local Similarity 100.0%; Pred.No.33;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 16 TCATTATTGATTTCAATTGTGCC 40  
|||||  
Db 170 TCATTATTGATTTCAATTGTGCC 194  
|||||

RESULT 13  
HSA325746/c  
LOCUS HSA325746 1090 bp DNA linear PRI 18-JUL-2002  
DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone  
NEI-850S.  
ACCESSION AJ325746  
VERSION AJ325746.1 GI:15870140  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1090)

AUTHORS Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvaasha,S.M., Podewski,R.M., Matsushkin,Y.G., Gyanchandani,A., Muravenko,O.V., Levitsky,V.G., Kolchanov,N.A., Protodopov,A.I., Kashuba,V.I., Kiselev,I.L., Wasserman,W., Wahlstedt,C. and Zabarovsky,E.R.

TITLE NotI flanking sequences: a tool for gene discovery and verification of the human genome

JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)

MEDLINE 22131767

PUBMED 12136098

REFERENCE 2 (bases 1 to 1090)

AUTHORS Zabarovsky,E.R.

TITLE Direct Submission

JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden

FEATURES  
source  
1. .1090  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 285 a 265 c 229 g 311 t

ORIGIN

Query Match 62.5%; Score 25; DB 9; Length 1090;  
Best Local Similarity 100.0%; Pred.No.33;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 16 TCATTATTGATTTCAATTGTGCC 40  
|||||  
Db 177 TCATTATTGATTTCAATTGTGCC 153  
|||||

RESULT 14  
AFI21976  
LOCUS AFI21976 1421 bp DNA linear ROD 07-DEC-1999  
DEFINITION Mus musculus odorant receptor S19 gene, complete cds.  
ACCESSION AFI21976  
VERSION AFI21976.2 GI:6532000  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 442 to 1421)

AUTHORS Malnic,B., Hirono,J., Sato,T. and Buck,L.B.

TITLE Combinatorial receptor codes for odors

JOURNAL Cell 96 (5), 713-723 (1999)

MEDLINE 99189756

PUBMED 10089886

REFERENCE 2 (bases 442 to 1421)

AUTHORS Malnic,B., Hirono,J., Sato,T. and Buck,L.B.

TITLE Direct Submission

JOURNAL Submitted (20-JAN-1999) Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, MA 02115, USA

REFERENCE 3 (bases 1 to 1421)

AUTHORS Malnic,B., Hirono,J., Sato,T. and Buck,L.B.

TITLE Direct Submission

JOURNAL Submitted (07-DEC-1999) Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, MA 02115, USA

REMARK Sequence update by submitter

COMMENT On Dec 7, 1999 this sequence version replaced gi:4680261.

FEATURES  
source  
1. .1421  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
4291..14310  
/product="odorant receptor S19"  
291..1310  
/note="G-protein-coupled receptor"  
/codon\_start=1  
/product="odorant receptor S19"  
/protein\_id="AAD27596.2"  
/db\_xref="GI:6532001"

mRNA

CDS

BASE COUNT 341 a 368 c 278 g 434 t

ORIGIN

Query Match 62.5%; Score 25; DB 10; Length 1421;  
Best Local Similarity 100.0%; Pred.No.33;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 16 TCATTATTGATTTCAATTGTGCC 40  
|||||  
Db 216 TCATTATTGATTTCAATTGTGCC 240  
|||||

```

RESULT 15
PGFRA/c
LOCUS      1423 bp      DNA      linear      INV 11-MAY-1992
DEFINITION P.polycephalum genomic DNA-fragment that binds Lupus type-Sm
            antibodies.
ACCESSION  X15815
VERSION    X15815.1 GI:3194
KEYWORDS   antibody-binding.
SOURCE     Physarum polycephalum (slime mold)
ORGANISM   Physarum polycephalum
            Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physariida;
            Physarum.
REFERENCE  1
AUTHORS    Retter,M.W. and Adams,D.S.
TITLE      Sequence of a developmentally transcribed physarum genomic DNA
            fragment that binds Lupus type-Sm Antibodies
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 1423)
AUTHORS    Adams,D.S. and Retter,M.W.
TITLE      Direct Submission
JOURNAL    Submitted (13-JUL-1989) D.S. Adams and M.W. Retter, WORCESTER
            POLYTECHNIC INSTITUTE, DEPT. OF BIOLOGY BIOTECHNOLOGY, 100
            INSTITUTE ROAD, WORCESTER, MA 01609-2280 USA
FEATURES   Location/Qualifiers
            1..1423
                /organism="Physarum polycephalum"
                /mol_type="genomic DNA"
                /strain="Carolina"
                /db_xref="taxon:5791"
                /cell_type="microplasmidia"
                /clone_lib="physarum, Sm-binding, genomic"
                /dev_stages="diploid, vegetative"
            misc_feature 1..1423
                /note="DNA fragment that binds Lupus type-Sm antibodies"
BASE COUNT 500 a 279 c 241 g 403 t
ORIGIN
Query Match      62.5%; Score 25; DB 3; Length 1423;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      16 TCATTATTGATTTCATTTCGCC 40
        |||||
Db       168 TCATTATTGATTTCATTTCGCC 144

```

Search completed: December 3, 2003, 13:18:41  
 Job time : 1089.6 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:41 ; Search time 191.405 Seconds  
(without alignments)  
564.131 Million cell updates/sec

Title: US-10-082-772-2

Perfect score: 40

Sequence: 1 aaggaagcgccgctcattattgatttcattgttc 40

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	40	100.0	40	22 AAF61393	Expression vector
2	25	62.5	33	16 AAF79564	PCR primer MB101.
3	25	62.5	33	24 AAD42556	Bacteriophage lamb
4	25	62.5	201	14 AAQ50319	Detection probe 11
5	25	62.5	282	25 ACC44676	Murine rDNA PCR pr
6	25	62.5	282	25 AAT16636	Artificial plant c
7	25	62.5	1763	20 AAZ30709	Rat neuronal immcd
8	25	62.5	2408	17 AAT30737	Rat cryptdin 2 gen

9	25	62.5	3049	24 AEN81284	Beet promoter poly
10	25	62.5	3484	19 AAV34987	Arabidopsis nemato
11	25	62.5	3757	19 AAV44332	Murine mCP-6 zymo
12	25	62.5	3757	19 AAV42708	Mouse mast cell pr
13	25	62.5	4346	25 ACC44716	Plasmid pSV40193at
14	25	62.5	4346	25 AAT16615	Artificial plant c
15	25	62.5	4909	16 AAT18924	Plasmid pAI261. S
16	25	62.5	5349	19 AAV23239	T-DNA of pTTS24. SY
17	25	62.5	5611	19 AAV23242	Plasmid pLH48. SY
18	25	62.5	5641	22 AAF30800	Vector plasmid pJM
19	25	62.5	5670	22 AAF30801	Vector plasmid pJM
20	25	62.5	5826	22 AAF30798	Vector plasmid pDY
21	25	62.5	6071	22 AAF30799	Vector plasmid pDY
22	25	62.5	7652	14 AAQ45682	Sequencing vector
23	23.2	58.0	2597	23 ABL28050	Drosophila melanog
24	22.6	56.5	28313	22 AAL36829	Human musculoskele
25	22.6	56.5	28313	25 ABX59817	cDNA encoding nove
26	22.6	56.5	29228	22 AAL36833	Human musculoskele
27	22.6	56.5	29228	25 ABX59821	cDNA encoding nove
28	22.4	56.0	450	24 ABK78409	Bacillus clausii g
29	22	55.0	34	25 ACC44615	Lambda integrase p
30	22	55.0	38	13 AAQ20343	PCR primer for amp
31	22	55.0	38	13 AAQ31423	33F2 toxin gene re
32	22	55.0	38	13 AAQ30955	Reverse PCR primer
33	22	55.0	38	13 AAQ32176	Reverse PCR primer
34	22	55.0	38	16 AAQ81165	B.t. toxin gene PC
35	22	55.0	38	18 AAT66807	Bacillus thuringie
36	22	55.0	38	18 AAT60053	Probe 63B/33F2 for
37	22	55.0	38	19 AAV59000	B.t. toxin gene pr
38	22	55.0	38	21 AAA65105	Probe #8 used to i
39	22	55.0	38	22 AAF28507	33F2 reverse prime
40	22	55.0	38	24 AAS16681	Bacillus thuringie
41	22	55.0	549	23 ABV56662	Human prostate exp
42	22	55.0	610	22 AAF79770	Bacteriophage lamb
43	22	55.0	1071	25 ACC44648	Lambda integrase E
44	22	55.0	1071	25 AET16637	Artificial plant c
45	22	55.0	5855	25 ACC44715	Plasmid pCXLamIntr

ALIGNMENTS

- RESULT 1
- ID AAF61393 standard; DNA; 40 BP.
- XX AAF61393;
- XX AC
- XX DT
- XX 05-JUN-2001 (first entry)
- XX Expression vector pKEX-2-XR Int gene PCR primer 3289.
- XX Sequence-specific recombination; SSR; integrase; gene therapy; somatic;
- XX targeted integration; Int gene; PCR primer; ss.
- XX Synthetic.
- XX DE19941186-A1.
- XX PD 01-MAR-2001.
- XX 30-AUG-1999; 99DE-1041186.
- XX 30-AUG-1999; 99DE-1041186.
- XX (DROE/) DROEGE P.
- XX Droege P;
- XX WPI; 2001-246016/26.
- XX Sequence-specific recombination of DNA in eukaryotes, useful
- XX particularly for somatic cell gene therapy, uses an integrase to effect



PT recombination between att sites -  
XX  
XX Example 1; Page 8; 24pp; German.  
XX  
XX This invention describes a novel sequence-specific recombination (SSR) of  
CC DNA in a eukaryotic cell, comprising introducing two DNA sequences (I,  
CC II) into a cell, using an integrase (Int) to effect SSR. The invention  
CC also describes (1) a nucleic acid comprising a 243 base pair sequence  
CC (III), fully defined in the specification, or its derivatives; and (2)  
CC vector containing (III), or its derivatives, plus a therapeutic gene, or  
CC its derivatives. The method is particularly used in somatic gene therapy  
CC in humans and animals, but can be applied more generally for gene  
CC transfer to animal or plant cells. The method is simple and controllable,  
CC and provides stable and targeted integration of selected DNA sequences.  
XX  
SQ Sequence 40 BP; 10 A; 8 C; 8 G; 14 T; 0 other;  
  
Query Match 100.0%; Score 40; DB 22; Length 40;  
Best Local Similarity 100.0%; Pred. No. 2.7e-06;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 AAGGAAAGCGCGCGCTCATTATTGATTTCATTTCATTTCGTC 40  
DB 1 AAGGAAAGCGCGCGCTCATTATTGATTTCATTTCATTTCGTC 40  
  
RESULT 2  
AAQ79564  
ID AAQ79564 standard; DNA; 33 BP.  
XX AAQ79564;  
AC  
XX 25-MAR-2003 (updated)  
DT 10-AUG-1995 (first entry)  
XX  
XX PCR primer MBL101.  
XX  
XX Thermostable DNA polymerase; PCR primer; ss.  
XX  
XX Synthetic.  
XX  
XX WO9426766-A1.  
XX  
XX 24-NOV-1994.  
XX  
XX 22-FEB-1994; 94WO-US01867.  
XX  
XX 19-FEB-1993; 93US-0021623.  
XX 22-FEB-1994; 94US-0202032.  
XX  
XX (BARN/) BARNES W M.  
XX  
XX Barnes WM;  
XX  
XX WPI; 1995-006692/01.  
XX  
XX DNA polymerase and formulations comprising it - allowing the  
XX amplification of sequences up to 35 kilobases and reducing the  
XX mutagenicity generated by the PCR process.  
XX  
XX Example; Page 52; 79pp; English.  
XX  
XX The example demonstrates PCR amplification using  
XX thermostable DNA polymerase KlenTaq-278.  
XX The primers used are Lc5 and Lc3. The template was pLc  
XX (a clone of an R colour control cDNA in maize).  
XX (Updated on 25-MAR-2003 to correct PN field.)  
XX (Updated on 25-MAR-2003 to correct PR field.)  
XX  
XX Sequence 33 BP; 7 A; 9 C; 2 G; 15 T; 0 other;  
  
Query Match 62.5%; Score 25; DB 16; Length 33;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
PT DNA polymerase and formulations comprising it - allowing the  
PT amplification of sequences up to 35 kilobases and reducing the  
PT mutagenicity generated by the PCR process.  
XX  
XX Example; Page 52; 79pp; English.  
XX  
XX The example demonstrates PCR amplification using  
XX thermostable DNA polymerase KlenTaq-278.  
XX The primers used are Lc5 and Lc3. The template was pLc  
XX (a clone of an R colour control cDNA in maize).  
XX (Updated on 25-MAR-2003 to correct PN field.)  
XX (Updated on 25-MAR-2003 to correct PR field.)  
XX  
XX Sequence 33 BP; 7 A; 9 C; 2 G; 15 T; 0 other;  
  
Query Match 62.5%; Score 25; DB 24; Length 33;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 16 TCATTATTGATTTCATTTCATTTCGTC 40  
DB 2 TCATTATTGATTTCATTTCATTTCGTC 26  
  
RESULT 3  
AAD42556  
ID AAD42556 standard; DNA; 33 BP.  
XX AAD42556;  
AC  
XX 15-NOV-2002 (first entry)  
XX  
XX Bacteriophage lambda clone MBL101 specific oligonucleotide.  
XX  
XX Enzymatic extension; DNA polymerase; 3'-exonuclease activity; ss.  
XX Bacteriophage lambda.  
XX  
XX US6410277-B1.  
XX 25-JUN-2002.  
XX  
XX 16-SEP-1997; 97US-0931818.  
XX  
XX 19-FEB-1993; 93US-0021623.  
XX 07-JUN-1995; 95US-0483535.  
XX  
XX (TAKI ) TAKARA SHUZO CO LTD.  
XX  
XX Barnes WM;  
XX  
XX WPI; 2002-588888/63.  
XX  
XX Improved polymerase chain reaction method utilizing 2 thermostable DNA  
XX polymerases either having or lacking 3'-exonuclease activity and  
XX derived from Pyrococcus furiosus, Thermotoga maritima, Thermococcus  
XX litoralis and Thermus aquaticus -  
XX  
XX Disclosure; Column 51; 37pp; English.  
XX  
XX The invention relates to an improved method for enzymatic extension of  
XX DNA strands when amplifying nucleic acid sequences by polymerase chain  
XX reaction (PCR), in which primer extension is catalysed using a  
XX thermostable DNA polymerase which lacks 3'-exonuclease activity and a  
XX thermostable DNA polymerase exhibiting 3'-exonuclease activity. The  
XX formulation of DNA polymerases is used in polymerase chain reactions  
XX (PCR). The present sequence is Bacteriophage lambda clone MBL101  
XX specific oligonucleotide.  
XX  
XX Sequence 33 BP; 7 A; 9 C; 2 G; 15 T; 0 other;  
  
Query Match 62.5%; Score 25; DB 24; Length 33;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 16 TCATTATTGATTTCATTTCATTTCGTC 40  
DB 2 TCATTATTGATTTCATTTCATTTCGTC 26  
  
RESULT 4  
AAQ50319  
ID AAQ50319 standard; DNA; 201 BP.  
XX AAQ50319;  
AC  
XX  
XX 27-APR-1994 (first entry)  
XX  
XX Detection probe 11.  
XX

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 16 TCATTATTGATTTCATTTCATTTCGTC 40  
DB 2 TCATTATTGATTTCATTTCATTTCGTC 26  
  
RESULT 3  
AAD42556  
ID AAD42556 standard; DNA; 33 BP.  
XX AAD42556;  
AC  
XX 15-NOV-2002 (first entry)  
XX  
XX Bacteriophage lambda clone MBL101 specific oligonucleotide.  
XX  
XX Enzymatic extension; DNA polymerase; 3'-exonuclease activity; ss.  
XX Bacteriophage lambda.  
XX  
XX US6410277-B1.  
XX 25-JUN-2002.  
XX  
XX 16-SEP-1997; 97US-0931818.  
XX  
XX 19-FEB-1993; 93US-0021623.  
XX 07-JUN-1995; 95US-0483535.  
XX  
XX (TAKI ) TAKARA SHUZO CO LTD.  
XX  
XX Barnes WM;  
XX  
XX WPI; 2002-588888/63.  
XX  
XX Improved polymerase chain reaction method utilizing 2 thermostable DNA  
XX polymerases either having or lacking 3'-exonuclease activity and  
XX derived from Pyrococcus furiosus, Thermotoga maritima, Thermococcus  
XX litoralis and Thermus aquaticus -  
XX  
XX Disclosure; Column 51; 37pp; English.  
XX  
XX The invention relates to an improved method for enzymatic extension of  
XX DNA strands when amplifying nucleic acid sequences by polymerase chain  
XX reaction (PCR), in which primer extension is catalysed using a  
XX thermostable DNA polymerase which lacks 3'-exonuclease activity and a  
XX thermostable DNA polymerase exhibiting 3'-exonuclease activity. The  
XX formulation of DNA polymerases is used in polymerase chain reactions  
XX (PCR). The present sequence is Bacteriophage lambda clone MBL101  
XX specific oligonucleotide.  
XX  
XX Sequence 33 BP; 7 A; 9 C; 2 G; 15 T; 0 other;  
  
Query Match 62.5%; Score 25; DB 24; Length 33;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 16 TCATTATTGATTTCATTTCATTTCGTC 40  
DB 2 TCATTATTGATTTCATTTCATTTCGTC 26  
  
RESULT 4  
AAQ50319  
ID AAQ50319 standard; DNA; 201 BP.  
XX AAQ50319;  
AC  
XX  
XX 27-APR-1994 (first entry)  
XX  
XX Detection probe 11.  
XX

KW	Probe; chip; cell; detection; label; capture probe; ss.
XX	Synthetic.
OS	JP05236997-A.
XX	17-SEP-1993.
PN	28-FEB-1992; 92JP-0042829.
PD	28-FEB-1992; 92JP-0042829.
XX	(HITA) HITACHI LTD.
PA	WIPI; 1993-330600/42.
XX	Chip for capturing polynucleotide - has several different complementary probes fixed on cells at different sites on single tip
PT	Disclosure; Page 8; 10pp; Japanese.
Pt	The sequences given in AAQ50309-26 are probes which were used on a chip for capturing a number of target polynucleotides. A group of these probes may be attached on cells which form part of a single chip. The oligonucleotides to be detected are labelled and are complementary to the set of capture probes used. The detection chip may be used to detect a number of different polynucleotides concurrently.
CC	Sequence 201 BP; 58 A; 38 C; 27 G; 78 T; 0 other;
CC	Query Match 62.5%; Score 25; DB 14; Length 201;
CC	Best Local Similarity 100.0%; Pred.No.2.6;
CC	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC	16 TCATTATTGATTTCAATTGTTC 40
CC	130 TCATTATTGATTTCAATTGTTC 154
CC	RESULT 5
CC	ACC44676
CC	ID ACC44676 standard; DNA; 282 BP.
CC	AC 44676;
CC	29-MAY-2003 (first entry)
CC	Murine rDNA PCR primer SEQ ID NO:72.
CC	Chromosome-based platform; artificial chromosome; eukaryotic chromosome; att site; integrase; recombinase; ACes; gene therapy; transgenic animal; Platform artificial chromosome expression system; gene; ds.
CC	Mus musculus.
CC	Synthetic.
CC	WO200297059-A2.
CC	05-DEC-2002.
CC	30-MAY-2002; 2002WO-US17452.
CC	30-MAY-2001; 2001US-294758P.
CC	21-MAR-2002; 2002US-366891P.
CC	(CHRO-) CHROMOS MOLECULAR SYSTEMS INC.
CC	Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E; Stewart S, Sheldard J;
CC	WIPI; 2003-140461/13.
CC	Producing artificial chromosome by introducing a nucleic acid into
CC	Novel eukaryotic chromosome comprising one or many att sites which permits site-directed integration in the presence of lambda-integrase, useful for site-specific recombination-directed integration of DNA of interest
CC	Example 2; Page 207; 272pp; English.
CC	The present invention describes a eukaryotic chromosome (I) comprising one or several att sites, where an att site is heterologous to the chromosome, and permits site-directed integration in the presence of lambda-integrase. Also described: (1) a platform artificial chromosome expression system (Acce); (ii) comprising several sites that participate in recombinase catalysed recombination; and (2) a method (M1) for introducing a heterologous nucleic acid into a platform artificial chromosome. (I) can be used in gene therapy. (M1) is useful for introducing a heterologous nucleic acid molecule into a platform artificial chromosome, preferably an Acce. (II) is useful for producing a transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or mammal) by introducing (ii) by cell fusion, lipid-mediated transfection, by a carrier system, microinjection, microcell fusion, electroporation, microparticle bombardment or direct DNA transfer into an embryonic cell, preferably a stem cell or an embryo. (II) comprises a heterologous nucleic acid that encodes a therapeutic product which is useful for making a library of Acce comprising random portions of a genome. ACC44612 to ACC44732 and ABP96650 to ABP96657 represent sequences used in the exemplification of the present invention.
CC	Sequence 282 BP; 79 A; 49 C; 39 G; 115 T; 0 other;
CC	Query Match 62.5%; Score 25; DB 25; Length 282;
CC	Best Local Similarity 100.0%; Pred.No.2.7;
CC	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC	16 TCATTATTGATTTCAATTGTTC 40
CC	243 TCATTATTGATTTCAATTGTTC 267
CC	RESULT 6
CC	ABTI16636
CC	ID ABTI16636 standard; DNA; 282 BP.
CC	AC ABTI16636;
CC	03-APR-2003 (first entry)
CC	Artificial plant chromosome related oligo SEQ ID No 48.
CC	Plant artificial chromosome; PAC; transgenic plant; vaccine; blood factor; herbicide; stress; agronomical; nutrient quality;
CC	bacterial artificial chromosome; BAC; yeast artificial chromosome; YAC; ds.
CC	Unidentified.
CC	WO200296923-A1.
CC	05-DEC-2002.
CC	30-MAY-2002; 2002WO-US17451.
CC	30-MAY-2001; 2001US-294687P.
CC	04-JUN-2001; 2001US-296329P.
CC	(CHRO-) CHROMOS MOLECULAR SYSTEMS INC.
CC	(AGRI-) AGRISOMA INC.
CC	Perez C, Fajianski SF, Perkins E;
CC	WIPI; 2003-140436/13.
CC	Producing artificial chromosome by introducing a nucleic acid into

PT plant cell, selecting artificial chromosome that has one or more repeat  
 PT regions with equivalent amounts of euchromatic and heterochromatic  
 PT nucleic acids  
 XX Disclosure; Page 263-264; 269pp; English.  
 XX  
 CC The invention relates to a novel method for producing plant artificial  
 CC chromosomes. The invention also relates to methods for targeting  
 CC insertion of heterologous DNA into plant artificial chromosomes, methods  
 CC for delivery of plant chromosomes to selected cells and tissues. The  
 CC isolated plant artificial chromosome (PAC) is useful for producing a  
 CC transgenic plant, which involves introducing the PAC into a plant cell.  
 CC The PAC comprises a heterologous nucleic acid encoding a gene product  
 CC such as enzymes, antisense RNA, tRNA, rDNA, structural proteins, marker  
 CC proteins, ligands, receptors, ribozymes, therapeutic proteins, and  
 CC biopharmaceutical proteins, vaccines, blood factors, antigens, hormones,  
 CC cytokines, growth factors, antibodies, or a product that provides for  
 CC resistance to diseases, insects, herbicides, or stress in a plant. The  
 CC heterologous nucleic acid optionally encodes a product that provides an  
 CC agronomically important trait in the plant, e.g. a product that alters  
 CC nutrient use and/or improves the nutrient quality of the plant. The  
 CC heterologous nucleic acid is contained within a bacterial artificial  
 CC chromosome (BAC) or a yeast artificial chromosome (YAC). This  
 CC polynucleotide sequence represents an oligo relating to the method for  
 CC producing plant artificial chromosomes of the invention.  
 XX  
 SQ Sequence 282 BP; 79 A; 49 C; 39 G; 115 T; 0 other;

Query Match 62.5%; Score 25; DB 25; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 2.7;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCATCAATTTGTCC 40  
 Db 243 TCATTATTGATTCATCAATTTGTCC 267

RESULT 7  
 AA230709/c  
 ID AA230709 standard; cDNA; 1763 BP.  
 XX  
 AC AA230709;  
 DT  
 DT 05-JAN-2000 (first entry)  
 XX  
 DE Rat neuronal immediate early gene cDNA clone R280.  
 XX  
 KW Immediate early gene; IEG; neuron; brain; function; growth factor;  
 KW transcription factor; signal transduction; cytoskeletal protein;  
 KW metabolic enzyme; learning; memory; synaptic transmission; tolerance;  
 KW neuronal plasticity; ds.

OS Rattus sp.  
 FH Key Location/Qualifiers  
 FT exon 1..1345  
 FT /\*tag= a  
 FT /codon\_start= 1174..1176  
 FT /note= "exon 1 codes for the 5' untranslated  
 FT intron region and cryptdin-1 prepro sequence"  
 FT 1345..1930  
 FT /\*tag= b  
 FT 1931..2408  
 FT /\*tag= c  
 FT /note= "exon 2 codes for cryptdin-2 and  
 FT 3' untranslated region"

XX WO9616075-A1.  
 PN (UJJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
 PA (BADI ) BASF-LYNX BIOSCIENCE AG.

XX  
 XX Worley PF, Lanahan A, Goetz B, Hiemisch H, Kuner R, Scheek S;  
 PI Nikolich K, Zhukovski E;  
 XX WPI; 1999-590697/50.

PT Novel genes and polypeptides, useful for treating conditions related to  
 PT a deficiency in nIEG responsiveness to a stimulus

XX Claim 1; Page 114-115; 134pp; English.  
 PS  
 XX This sequence represents rat neuronal immediate early gene (IEG) cDNA  
 CC clone R280. An IEG is a gene whose expression is rapidly  
 CC increased immediately following a stimulus e.g., neuronal stimulation.  
 CC Such neuronal IEGs have been found to encode a variety of proteins,  
 CC including transcription factors, cytoskeletal proteins, growth factors  
 CC and metabolic enzymes, as well as proteins involved in signal  
 CC transduction. The identification of neuronal IEGs and the proteins they  
 CC encode may provide important information about the function of neurons  
 CC in, for example, learning, memory, synaptic transmission, tolerance and  
 CC neuronal plasticity. Neuronal IEGs, neuronal IEG protein products, cells  
 CC expressing neuronal IEGs and antibodies against neuronal IEG proteins  
 CC can be used to treat an animal with a deficiency in neuronal IEG  
 CC responsibility to stimuli, such that the effect of the deficiency is  
 CC minimised. The deficiency may be a reduced or elevated level of  
 CC expression of an IEG. The neuronal stimulus comprises a maximal  
 CC electroconvulsive seizure and its effects influence learning or memory.  
 CC The IEGs and protein products are useful in identifying compounds  
 CC that modulate the expression or activity of IEG nucleic acids or  
 CC proteins, respectively.  
 XX  
 SQ Sequence 1763 BP; 478 A; 402 C; 452 G; 424 T; 7 other;

Query Match 62.5%; Score 25; DB 20; Length 1763;  
 Best Local Similarity 100.0%; Pred. No. 3.5;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCATCAATTTGTCC 40  
 Db 203 TCATTATTGATTCATCAATTTGTCC 179

RESULT 8  
 AAT30737  
 ID AAT30737 standard; DNA; 2408 BP.  
 XX  
 AC AAT30737;  
 DT  
 DT 20-OCT-1996 (first entry)  
 XX  
 DE Rat cryptdin 2 gene.  
 XX  
 KW Cryptdin 2; antibiotic; antimicrobial; defensin; inflammation;  
 KW antiinflammatory; inflammatory bowel disease; pancreatitis;  
 KW cancer; tumour; ileitis; ds.

OS Rattus sp.  
 FH Key Location/Qualifiers  
 FT exon 1..1345  
 FT /\*tag= a  
 FT /codon\_start= 1174..1176  
 FT /note= "exon 1 codes for the 5' untranslated  
 FT intron region and cryptdin-1 prepro sequence"  
 FT 1345..1930  
 FT /\*tag= b  
 FT 1931..2408  
 FT /\*tag= c  
 FT /note= "exon 2 codes for cryptdin-2 and  
 FT 3' untranslated region"

XX WO9616075-A1.  
 PN 30-MAY-1996.  
 PD  
 XX 05-OCT-1995; 95WO-US13328.  
 PF  
 XX 18-NOV-1994; 94US-0342268.  
 PR  
 XX (SHRI-) SHRINER'S HOSPITAL FOR CRIPPLED CHILDREN.  
 PA (REGC ) UNIV CALIFORNIA.

PT roots or aerial parts -  
XX  
PS Claim 1; Page 30-31; 57pp; German.  
XX  
CC The invention relates to a promoter (A) that is: (i) any of the sequences  
CC ABN81281-ABN81284; (ii) the complement of (i); or (iii) a sequence that  
CC hybridises to (ii). (A) are used, specifically in Beta vulgaris, to  
CC control expression of transgenes, particularly to alter carbohydrate  
CC metabolism; reduce loss of storage substances; express invertase  
CC inhibitor, fructosyl transferase, levan sucrase or genes that encode  
CC transporter proteins for nitrogen compounds, or increase resistance to,  
CC or tolerance of, pathogens. (A) provide tissue-specific transgene  
CC expression, either in roots or above-ground parts, so avoid pleiotropic  
CC effects, e.g. when expressing invertase inhibitor.  
XX  
SQ Sequence 3049 BP; 783 A; 699 C; 876 G; 691 T; 0 other;  
Query Match 62.5%; Score 25; DB 24; Length 3049;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 16 TCATTATTGATTCAATTTGTGCC 40  
DB 2485 TCATTATTGATTCAATTTGTGCC 2509  
RESULT 10  
AAV34987  
ID AAV34987 standard; DNA; 3484 BP.  
XX  
AC AAV34987;  
XX  
DT 28-SEP-1998 (first entry)  
XX  
DE Arabidopsis nematode feeding site-preferential promoter.  
XX  
KW Promoter; nematode feeding site; root knot nematode; cyst nematode;  
KW Meloidogyne incognita; Heterodera schachtii; Globodera pallida;  
XX transgenic plant; pest resistance; crop protection; ds.  
OS Arabidopsis thaliana ecotype C24.  
XX  
FH Key Location/Qualifiers  
FT CDS 3482  
FT /\*tag= a  
FT /codon\_start= 3482...3484  
XX  
PN WO9822599-A1.  
XX  
PD 28-MAY-1998.  
XX  
PF 18-NOV-1997; 97WO-EP06472.  
XX  
PR 18-NOV-1996; 96EP-0203213.  
XX  
PA (MOGE-) MOGEN INT NV.  
XX  
PI Goddijn OJM, Klap J, Ohl SA, Sijmons PC, Van Der Lee FM;  
XX WPI; 1998-312484/27.  
XX  
PT New Arabidopsis thaliana DNA promoter - useful for, e.g. generating  
XX plant reduced susceptibility to plant parasitic nematodes  
PS Claim 1; Page 22-24; 47pp; English.  
XX  
CC This DNA fragment obtainable from Arabidopsis thaliana is capable of  
CC promoting root knot and cyst nematode-inducible transcription of an  
CC associated DNA sequence when reintroduced into a plant. The  
CC promoterless GUS construct binary vector pMOG553 was mobilised by  
CC triparental mating to Agrobacterium tumefaciens MOG101 and the  
CC resulting strain was used for Arabidopsis root transformation.  
CC Line pMOG553#25 was identified as a line which showed strong GUS

XX  
PI Ouelllette AJ, Seilsted MB;  
XX  
DR WPI; 1996-268527/27.  
XX  
P-PSDB; AAR98787, AAR98792.  
XX  
XX New isolated cryptidin peptide(s) - which have antimicrobial  
PT activity, used partic. in the detection and treatment of  
PT inflammatory pathologies  
XX  
PS Claim 31; Page 64-65; 103pp; English.  
XX  
CC The rat cryptidin-2 gene (T30737) codes for the precursor  
CC (R98787) of cryptidin-2 (R98792), an antimicrobial peptide  
CC that exhibits activity against a broad range of intestinal and  
CC opportunistic pathogens. Rat cryptidin-1, -2, and -3 genes (see  
CC also T30736 and T30738) were isolated from a Sprague-Dawley  
CC genomic library cloned in EMBL3. The cryptidin-2 gene, or  
CC corresponding cDNA (see also T30734), can be used to produce  
CC large amounts of cryptidin-2 for use in treating inflammatory  
CC pathologies of the intestines.  
XX  
SQ Sequence 2408 BP; 640 A; 531 C; 557 G; 680 T; 0 other;  
Query Match 62.5%; Score 25; DB 17; Length 2408;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 16 TCATTATTGATTCAATTTGTGCC 40  
DB 495 TCATTATTGATTCAATTTGTGCC 519  
RESULT 9  
ABN81284  
ID ABN81284 standard; DNA; 3049 BP.  
XX  
AC ABN81284;  
XX  
DT 22-AUG-2002 (first entry)  
XX  
DE Beet promoter polynucleotide SEQ ID NO 4.  
XX  
KW Beet; promoter; carbohydrate metabolism; invertase inhibitor;  
KW fructosyl transferase; levan sucrase; nitrogen transporter protein;  
KW pathogen resistance; plant; transgenic; ds.  
XX  
OS Beta vulgaris.  
XX  
FH Key Location/Qualifiers  
FT Promoter 1..2998  
FT /\*tag= a  
FT mRNA 2928..3049  
FT /\*tag= b  
FT TATA\_signal 2877..2883  
FT /\*tag= c  
XX  
EP1207204-A1.  
XX  
PN  
XX  
PD 22-MAY-2002.  
XX  
PF 16-NOV-2000; 2000EP-0124989.  
XX  
PR 16-NOV-2000; 2000EP-0124989.  
XX  
PA (KWSS-) KWS SAAT AG.  
XX  
PI Hehl R, Kloos D, Stahl DJ;  
XX  
DR WPI; 2002-437465/47.  
XX  
XX New tissue-specific promoters from Beta vulgaris, useful e.g. for  
PT altering carbohydrate metabolism, express transgenes selectively in

expression inside syncytia and giant cells induced by the cyst nematode *Heterodera schachtii* and the root knot nematode *Meloidogyne incognita*, respectively. Promoter tags from line pWGS53#25 were sequenced. The claimed DNA fragment is nematode feeding site-specific. Also claimed are: (1) a portion or variant of the above sequence capable of promoting root knot and cyst nematode inducible transcription of an associated DNA sequence when reintroduced into a plant; (2) a chimeric DNA sequence comprising a DNA fragment as above plus a DNA sequence which is not naturally under its transcriptional control, and which causes the production of a plant cell-disruptive substance, preferably barnase; (3) a replicon comprising the chimeric DNA sequence of (2), or the above DNA and at least 1 restriction endonuclease recognition site; (4) a microorganism containing the replicon of (3); (5) a plant cell having incorporated into its genome the chimeric DNA of (2); (6) a root system of a plant consisting of the cells of (5); (7) a plant consisting of the cells of (5), preferably a dicotyledonous plant, especially a potato plant; (8) a plant grafted onto the root system of (6); (9) a part of a plant, selected from seeds, flowers, tubers, roots, leaves, fruits, pollen and wood, obtained from the plant of (7) or (8), and (10) a crop consisting of the plants of (7) or (8). The DNA fragment can be used to identify subfragments capable of promoting transcription of an associated DNA sequence in a plant. It can also be used for making hybrid regulatory DNA sequences. The chimeric DNA sequence of (2) can be used for transforming plants (all claimed). The DNA sequence can be used to reduce the susceptibility of a plant to parasitic nematodes.

SQ Sequence 3484 BP; 1143 A; 583 C; 625 G; 1133 T; 0 other;

Query Match 62.5%; Score 25; DB 19; Length 3484;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTTCAATTGTGTC 40  
Db 507 TCATTATTGATTTCAATTGTGTC 531

RESULT 11

AAV44332

ID AAV44332 standard; DNA; 3757 BP.

AC AAV44332;

DT 24-NOV-1998 (first entry)

DE Murine mMCP-6 zymogen gene.

MCP-6 zymogen; mouse; mast cell protease 6; mMCP-7; tryptase-7; blood clot; anticoagulant; myocardial infarction; reocclusion; thromboembolism; cerebral embolism; thrombosis; therapy; ss.

OS Mus sp.

Key Location/Qualifiers

CDS 1803..3098

FT /\*tag= a

FT /note= "contains introns"

FT exon 1803..1872

FT /\*tag= b

FT /number= 1

FT intron 1873..2011

FT /\*tag= c

FT /number= 1

FT exon 2012..2177

FT /\*tag= d

FT /number= 2

FT intron 2178..2297

FT /\*tag= e

FT /number= 2

FT exon 2298..2563

FT /\*tag= f

FT intron /number= 3  
FT 2564..2696  
FT /\*tag= g  
FT /number= 3  
FT exon 2697..2860  
FT /\*tag= h  
FT intron /number= 4  
FT 2861..2933  
FT /\*tag= i  
FT /number= 4  
FT exon 2934..3098  
FT /\*tag= j  
FT /number= 5

WO9824386-A1.

XX 11-JUN-1998.

XX 25-NOV-1997; 97WO-US21620.

XX 04-DEC-1996; 96US-0032354.

PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.

XX Stevens RL;

XX WPI; 1998-333308/29.

DR P-PSDB; AAW64242.

XX New compositions containing tryptase-7, e.g. mouse mast cell protease-7 - are used to treat clot formation in e.g. myocardial infarction, reocclusion following angioplasty or pulmonary thromboembolism

PS Disclosure; Page 68-69; 92pp; English.

XX This nucleotide sequence includes a coding region for mouse mast cell protease 6 (mMCP-6, see AAW64242). The invention provides: CC compositions comprising an isolated tryptase-7 (such as mMCP-7, see AAW64233); a method for treating a blood clot by administering a CC nucleic acid molecule that codes for a tryptase-7, or an expression CC product; a nucleic acid encoding a serine protease (SP); and a CC method of producing a mature SP. Tryptase-7 polypeptides can be CC used for treating disorders mediated by undesirable thrombus clot CC formation such as myocardial infarction and reocclusion following CC angioplasty of blood clots associated with pulmonary CC thromboembolism, deep vein thrombosis, cerebral embolism, renal CC vein and peripheral arterial thrombosis.

SQ Sequence 3757 BP; 802 A; 1014 C; 1003 G; 938 T; 0 other;

Query Match 62.5%; Score 25; DB 19; Length 3757;

Best Local Similarity 100.0%; Pred. No. 3.8;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTTCAATTGTGTC 40

Db 461 TCATTATTGATTTCAATTGTGTC 485

RESULT 12

AAV42708

ID AAV42708 standard; cDNA; 3757 BP.

XX AAV42708;

XX 27-OCT-1998 (first entry)

XX Mouse mast cell protease (mMCP-6) nucleic acid sequence.

XX Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma; tryptase-6 protein; inflammatory disorder; allergic rhinitis; urticaria; antioedema; eczematous dermatitis; atopic dermatitis; anaphylaxis;

KW hyperproliferative skin disease; peptic ulcer; hyperresponsiveness;  
KW inflammatory skin condition; ss.  
XX  
OS Mus sp.  
XX  
PN WO9833812-A1.  
XX  
PD 06-AUG-1998.  
XX  
PF 30-JAN-1998; 98WO-US01865.  
XX  
PR 05-FEB-1997; 97US-0037090.  
XX  
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
XX  
PI Huang C, Stevens RL;  
XX  
DR WPI; 1998-437390/37.  
XX  
PT Trypsin-6 complex inhibitory peptides - used to treat mast  
PT cell-mediated inflammatory disorders e.g. asthma  
XX  
PS Disclosure; Pages 42-43; 69pp; English.  
XX  
CC This represents the cDNA sequence of the mouse mast cell protease  
CC (mMCP-6). The invention provides sequences shown in AA663160 to  
CC AA663169 that are inhibitors of mMCP-6. These peptides which are  
CC trypsin-6 complex inhibitors, can be used for treating a mast  
CC cell-mediated inflammatory disorder. The inhibitors can be used to treat  
CC inflammatory disorders including asthma, allergic rhinitis, urticaria and  
CC antioedema, eczematous dermatitis (atopic dermatitis), hyperproliferative  
CC skin disease, anaphylaxis, peptic ulcers, inflammatory bowel disorder,  
CC hyperresponsiveness and inflammatory skin conditions.  
XX  
SQ Sequence 3757 BP; 802 A; 1014 C; 1003 G; 938 T; 0 other;  
  
Query Match 62.5%; Score 25; DB 19; Length 3757;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 16 TCATTATTGATTTCAATTGTCC 40  
Db 461 TCATTATTGATTTCAATTGTCC 485  
  
RESULT 13  
ACC44716  
ID ACC44716 standard; DNA; 4346 BP.  
XX  
AC ACC44716;  
XX  
DT 29-MAY-2003 (first entry)  
XX  
DE Plasmid pSV40193attPeensepur nucleotide sequence SEQ ID NO:113.  
XX  
KW Chromosome-based platform; artificial chromosome; eukaryotic chromosome;  
KW att site; integrase; recombinase; ACes; gene therapy; transgenic animal;  
KW platform artificial chromosome expression system; Gene; ds.  
XX  
OS Bacteriophage lambda.  
OS Synthetic.  
XX  
PN WO200297059-A2.  
XX  
PD 05-DEC-2002.  
XX  
PF 30-MAY-2002; 2002WO-US17452.  
XX  
PR 30-MAY-2001; 2001US-294758P.  
PR 21-MAR-2002; 2002US-366891P.  
XX  
PA (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.  
XX

PI Perkins B, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E;  
PI Stewart S, Shellard J;  
XX  
DR WPI; 2003-140461/13.  
XX  
PT Novel eukaryotic chromosome comprising one or many att sites which  
PT permits site-directed integration in the presence of lambda-integrase,  
PT useful for site-specific recombination-directed integration of DNA of  
PT interest  
XX  
PS Example 3; Page 244-245; 272pp; English.  
XX  
CC The present invention describes a eukaryotic chromosome (I) comprising  
CC one or several att sites, where an att site is heterologous to the  
CC chromosome, and permits site-directed integration in the presence of  
CC lambda-integrase. Also described: (1) a platform artificial chromosome  
CC expression system (ACes) (II) comprising several sites that participate  
CC in recombinase catalysed recombination; and (2) a method (M1) for  
CC introducing a heterologous nucleic acid into a platform artificial  
CC chromosome. (I) can be used in gene therapy. (M1) is useful for  
CC introducing a heterologous nucleic acid molecule into a platform  
CC artificial chromosome, preferably an ACes. (II) is useful for producing a  
CC transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or  
CC mammal) by introducing (II) by cell fusion, lipid-mediated transfection,  
CC by a carrier system, microinjection, microcell fusion, electroporation,  
CC microprojectile bombardment or direct DNA transfer into an embryonic  
CC cell, preferably a stem cell or an embryo. (II) comprises a heterologous  
CC nucleic acid that encodes a therapeutic product which is useful for  
CC making a library of ACes comprising random portions of a genome. ACC44612  
CC to ACC44732 and ABP96650 to ABP96657 represent sequences used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 4346 BP; 1022 A; 1208 C; 1094 G; 1022 T; 0 other;  
  
Query Match 62.5%; Score 25; DB 25; Length 4346;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 16 TCATTATTGATTTCAATTGTCC 40  
Db 4298 TCATTATTGATTTCAATTGTCC 4322  
  
RESULT 14  
ABT16615  
ID ABT16615 standard; DNA; 4346 BP.  
XX  
AC ABT16615;  
XX  
DT 03-APR-2003 (first entry)  
XX  
DE Artificial plant chromosome related plasmid DNA SEQ ID NO 26.  
XX  
KW Plant artificial chromosome; PAC; transgenic plant; vaccine;  
KW blood factor; herbicide; stress; agronomical; nutrient quality;  
KW bacterial artificial chromosome; BAC; yeast artificial chromosome; YAC;  
KW ds.  
XX  
OS Unidentified.  
XX  
PN WO200296923-A1.  
XX  
PD 05-DEC-2002.  
XX  
PF 30-MAY-2002; 2002WO-US17451.  
XX  
PR 30-MAY-2001; 2001US-294687P.  
PR 04-JUN-2001; 2001US-296329P.  
XX  
XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.  
PA (AGRI-) AGRISOMA INC.  
XX  
PI Perez C, Rabinanski SF, Perkins E;

XX WPI; 2003-140436/13.

XX Producing artificial chromosome by introducing a nucleic acid into

PT plant cell, selecting artificial chromosome that has one or more repeat

PT regions with equivalent amounts of euchromatic and heterochromatic

PT nucleic acids -

XX Example 19; Page 255-256; 269pp; English.

XX The invention relates to a novel method for producing plant artificial

CC chromosomes. The invention also relates to methods for targeting

CC insertion of heterologous DNA into plant artificial chromosomes, methods

CC for delivery of plant chromosomes to selected cells and tissues. The

CC isolated plant artificial chromosome (PAC) is useful for producing a

CC transgenic plant, which involves introducing the PAC into a plant cell.

CC The PAC comprises a heterologous nucleic acid encoding a gene product

CC such as enzymes, antisense RNA, rDNA, structural proteins, marker

CC proteins, ligands, receptors, ribozymes, therapeutic proteins, and

CC biopharmaceutical proteins, vaccines, blood factors, antigens, hormones,

CC cytokines, growth factors, antibodies, or a product that provides for

CC resistance to diseases, insects, herbicides, or stress in a plant. The

CC heterologous nucleic acid optionally encodes a product that provides an

CC agronomically important trait in the plant, e.g. a product that alters

CC nutrient use and/or improves the nutrient quality of the plant. The

CC heterologous nucleic acid is contained within a bacterial artificial

CC chromosome (BAC) or a yeast artificial chromosome (YAC). This

CC polynucleotide sequence represents the DNA of a plasmid used in the

CC method of the invention.

XX Sequence 4346 BP; 1022 A; 1208 C; 1094 G; 1022 T; 0 other;

SQ

Query Match 62.5%; Score 25; DB 25; Length 4346;

Best Local Similarity 100.0%; Pred. No. 3.9;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCATTGTC 40

Db 4298 TCATTATTGATTTCATTGTC 4322

RESULT 15

AAT18924/C

ID AAT18924 standard; DNA; 4909 BP.

XX AC AAT18924;

XX 17-JAN-1997 (first entry)

XX Plasmid pA1261.

XX Spider; dragline protein; variant; monomer; polymer; circular;

KW fibre forming region; Spidroin 1; Nephila clavipes; Dp1; mimic;

KW DP-1A analogue; fibre; high tensile strength; elasticity; clothing;

KW rope; surgical suture; implant; reinforcement; film; coating; ss.

XX Synthetic.

OS

XX WO9429450-A2.

XX 22-DEC-1994.

XX 15-JUN-1994; 94WO-US06689.

XX 15-JUN-1993; 93US-0077600.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Fahnstock SR;

PI WPI; 1995-036479/05.

XX New synthetic variants of spider dragline protein - for making

PT

PT fibres useful as clothing, surgical silk, plastic reinforcement

PT etc., also related DNA, vectors and transformed cells

XX Example 1; Fig 13; 168pp; English.

XX This sequence represents the complete nucleotide sequence of the plasmid

CC PA1261. This plasmid was used in the construction of the vector pP510

CC which was used to express synthetic spider dragline variants, pp-1A.9

CC and DP-1B.9. PA1261 comprises a replication origin active in E. coli,

CC a selectable genetic marker which is a gene conferring resistance to

CC ampicillin, sites for the restriction endonucleases BamHI and BglII with

CC no essential sequences between them, and a third restriction site for

CC PstI, located within the selectable marker which produces cohesive ends

CC incompatible with those produced by BamHI and BglII. The polypeptide

CC monomers are variants based on a consensus sequence derived from the

CC fibre forming regions of spider dragline protein, esp. the natural

CC protein 1 (Spidroin 1) from Nephila clavipes. Synthetic analogues

CC of Dp1 were designed to mimic the repeating consensus sequence of the

CC natural protein and the pattern of variation among individual repeats.

CC DP-1A analogues are composed of a tandemly repeated 101 amino acid

CC monomer which comprises four repeats which differ from the consensus

CC sequence given in AA06201, according to the pattern (1)-(5):

CC (1) the poly-alanine sequence varies in length from 0-7

CC residues; (2) when the entire poly-alanine sequence is deleted,

CC so also is the surrounding sequence encompassing AGRGLGQAGAGG;

CC (3) aside from the poly-alanine sequence, deletions usually

CC encompass integral multiples of three consecutive residues;

CC (4) deletion of GYG is generally accompanied by deletion of GRG

CC in the same sequence; and

CC (5) a repeat in which the entire poly-alanine sequence is

CC deleted is generally preceded by a repeat containing six alanine

CC residues.

CC The proteins may be used to produce fibres of high tensile strength and

CC elasticity, suitable for clothing, rope, surgical sutures, biomaterials

CC for implants, plastic reinforcements, films, coatings, etc.

XX

SQ Sequence 4909 BP; 1216 A; 1251 C; 1258 G; 1184 T; 0 other;

Query Match 62.5%; Score 25; DB 16; Length 4909;

Best Local Similarity 100.0%; Pred. No. 4;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCATTGTC 40

Db 2288 TCATTATTGATTTCATTGTC 2264

Search completed: December 3, 2003, 12:23:35

Job time : 192.405 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 47.6033 Seconds  
(without alignments)  
370.884 Million cell updates/sec

Title: US-10-082-772-2

Perfect score: 40

Sequence: 1 aaggaagcggcgctcatttggatttgcatttgc 40

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/prodata/1/ina/5A COMB.seq:\*

2: /cgn2\_6/prodata/1/ina/5B COMB.seq:\*

3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*

4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*

5: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq:\*

6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	62.5	33	US-09-011-223C-5	Sequence 5, Appli
2	25	62.5	201	US-08-021-667A-18	Sequence 18, Appl
3	25	62.5	201	US-08-410-544-18	Sequence 18, Appl
4	25	62.5	201	US-08-728-785A-18	Sequence 18, Appl
5	25	62.5	2408	US-08-486-013-69	Sequence 69, Appl
6	25	62.5	2408	US-08-482-279-69	Sequence 69, Appl
7	25	62.5	2408	US-08-342-268-69	Sequence 69, Appl
8	25	62.5	2408	US-09-015-968-69	Sequence 69, Appl
9	25	62.5	2408	US-09-397-386-69	Sequence 69, Appl
10	25	62.5	3484	US-09-308-090-1	Sequence 1, Appli
11	25	62.5	3484	US-09-380-090A-1	Sequence 1, Appli
12	25	62.5	3757	US-09-016-366A-13	Sequence 13, Appl
13	25	62.5	3757	US-08-978-404B-19	Sequence 19, Appl
14	25	62.5	4909	US-08-556-978B-78	Sequence 78, Appl
15	25	62.5	5349	US-09-068-101-7	Sequence 7, Appli
16	25	62.5	5611	US-09-068-101-10	Sequence 10, Appl
17	25	62.5	7652	US-07-590-988A-1	Sequence 1, Appli
18	25	62.5	38584	US-09-453-702B-50	Sequence 50, Appl
19	22	55.0	38	US-08-049-783-14	Sequence 14, Appl
20	22	55.0	38	US-08-158-232-28	Sequence 28, Appl
21	22	55.0	38	US-08-304-626-28	Sequence 28, Appl
22	22	55.0	38	US-08-316-301A-35	Sequence 35, Appl
23	22	55.0	38	US-08-611-928-28	Sequence 28, Appl
24	22	55.0	38	US-09-224-024-12	Sequence 12, Appl
25	22	55.0	38	US-09-173-891-28	Sequence 28, Appl
26	22	55.0	38	US-09-076-137-35	Sequence 35, Appl
27	22	55.0	38	PCT-US92-03624-35	Sequence 35, Appl

28	22	55.0	38	5	PCT-US94-07902-12	Sequence 12, Appl
29	22	55.0	219	4	US-09-107-532A-31	Sequence 31, Appl
30	21.8	54.5	9827	4	US-09-453-702B-66	Sequence 66, Appl
31	21.4	53.5	1132	3	US-09-036-574-8	Sequence 8, Appl
32	20.4	51.0	3561	1	US-08-485-568A-3	Sequence 3, Appl
33	20.4	51.0	3561	1	US-08-357-698-5	Sequence 5, Appl
34	20.4	51.0	3561	2	US-08-590-554A-3	Sequence 3, Appl
35	20.4	51.0	3561	2	US-09-184-223-3	Sequence 3, Appl
36	20.4	51.0	3561	5	PCT-US93-12882-5	Sequence 5, Appl
37	20.4	51.0	4281	4	US-09-357-206A-8	Sequence 8, Appl
38	20.4	51.0	10881	4	US-09-357-206A-9	Sequence 9, Appl
39	20.4	51.0	12886	4	US-09-357-206A-1	Sequence 1, Appl
40	20.2	50.5	2377	2	US-08-967-101-26	Sequence 26, Appl
41	20.2	50.5	2377	3	US-08-592-541-26	Sequence 26, Appl
42	20.2	50.5	2377	3	US-09-124-698-26	Sequence 26, Appl
43	20.2	50.5	2377	3	US-09-127-480-26	Sequence 26, Appl
44	20.2	50.5	2377	3	US-08-496-841C-26	Sequence 26, Appl
45	20.2	50.5	2377	4	US-09-124-523-26	Sequence 26, Appl

## ALIGNMENTS

### RESULT 1

US-09-011-223C-5  
; Sequence 5, Application US/09011223C  
; Patent No. 6255048  
; GENERAL INFORMATION:  
; APPLICANT: HOSOI, Shigeru  
; KADUCHI, Sachiko  
; KAJIMA, Makiko  
; TITLE OF INVENTION: HIGH SENSITIVITY FLUOROIMMUNOASSAY  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: PILLSBURY, MADISON & SUTRO  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOCS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/011,223C  
; FILING DATE: 09-Feb-1998  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul E. White, Jr.  
; REGISTRATION NUMBER: 32,011  
; REFERENCE/DOCKET NUMBER: 7898/244916  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627CUSH  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 33  
; TYPE: Nucleic acid  
; STRANDEDNESS: Single-stranded  
; TOPOLOGY: Linear  
; MOLECULE TYPE: DNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-011-223C-5

Query Match 62.5%; Score 25; DB 3; Length 33;  
Best Local Similarity 100.0%; Pred. NO. 0.28;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

16 TCATTATTGATTTCATTGTC 40



Db 2 TCATTATTGATTCAATTGTGCC 26

RESULT 2  
US-08-021-667A-18  
; Sequence 18, Application US/08021667A  
; Patent No. 5434049  
; GENERAL INFORMATION:  
; APPLICANT: Okano, Kazunori  
; APPLICANT: Kambara, Hideki  
; TITLE OF INVENTION: POLYNUCLEOTIDE CAPTURING TIP AND  
; TITLE OF INVENTION: POLYNUCLEOTIDE PREPARATIVE METHOD AND DETECTION  
; TITLE OF INVENTION: METHOD USING SAME  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Antonelli, Terry, Stout & Kraus  
; STREET: Suite 600, 1919 Pennsylvania Ave., NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/021,667A  
; FILING DATE: 19930224  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Terry, David T.  
; REGISTRATION NUMBER: 20,178  
; REFERENCE/DOCKET NUMBER: 520.31930X00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-828-0300  
; TELEFAX: 202-828-0380  
; TELEX: 248545  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 201 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: YES  
; ANTI-SENSE: NO  
US-08-021-667A-18

Query Match 62.5%; Score 25; DB 1; Length 201;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCAATTGTGCC 40  
|||||  
Db 130 TCATTATTGATTCAATTGTGCC 154

RESULT 3  
US-08-410-544-18  
; Sequence 18, Application US/08410544  
; Patent No. 5607646  
; GENERAL INFORMATION:  
; APPLICANT: Okano, Kazunori  
; APPLICANT: Kambara, Hideki  
; TITLE OF INVENTION: POLYNUCLEOTIDE CAPTURING TIP AND  
; TITLE OF INVENTION: POLYNUCLEOTIDE PREPARATIVE METHOD AND DETECTION  
; TITLE OF INVENTION: METHOD USING SAME  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Antonelli, Terry, Stout & Kraus  
; STREET: Suite 600, 1919 Pennsylvania Ave., NW  
; CITY: Washington

Qy 16 TCATTATTGATTCAATTGTGCC 40  
|||||  
Db 130 TCATTATTGATTCAATTGTGCC 154

RESULT 4  
US-08-728-785A-18  
; Sequence 18, Application US/08728785A  
; Patent No. 5817506  
; GENERAL INFORMATION:  
; APPLICANT: Okano, Kazunori  
; APPLICANT: Kambara, Hideki  
; TITLE OF INVENTION: POLYNUCLEOTIDE CAPTURING TIP AND  
; TITLE OF INVENTION: POLYNUCLEOTIDE PREPARATIVE METHOD AND DETECTION  
; TITLE OF INVENTION: METHOD USING SAME  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Antonelli, Terry, Stout & Kraus  
; STREET: Suite 1800, 1300 No. 5817506th Seventeenth St.  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22209  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/728,785A  
; FILING DATE: 10-OCT-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/410,544  
; FILING DATE: 21-MAR-1995

Qy 16 TCATTATTGATTCAATTGTGCC 40  
|||||  
Db 130 TCATTATTGATTCAATTGTGCC 154

```

/ LENGTH: 2408 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-486-013-69

Query Match          62.5%; Score 25; DB 1; Length 2408;
Best Local Similarity 100.0%; Pred.No. 0.56;
Matches 25; Conservative 0; Mismatches 0; Indels 0;

QY      16 TCATTATTGATTCAATTTGTCC 40
      |||||||
DB      495 TCATTATTGATTCAATTTGTCC 519

RESULT 6
US-08-482-279-69
/ Sequence 69, Application US/08482279
/ Patent No. 5840498
/ GENERAL INFORMATION:
/ APPLICANT: Selsted, Michael E.
/ APPLICANT: Ouellette, Andre J.
/ TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods
/ TITLE OF INVENTION: Of Their Use
/ NUMBER OF SEQUENCES: 70
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Campbell and Flores
/ STREET: 4370 La Jolla Village Drive, Suite 700
/ CITY: San Diego
/ STATE: California
/ COUNTRY: USA
/ ZIP: 92122
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/482,279
/ FILING DATE:
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/342,268
/ FILING DATE: 18-NOV-1994
/ APPLICATION NUMBER: US 07/930,649
/ FILING DATE: 14-AUG-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/889,020
/ FILING DATE: 26-MAY-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Campbell, Cathryn A.
/ REGISTRATION NUMBER: 31,815
/ REFERENCE/POCKET NUMBER: P-UC 1206
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (619) 535-9001
/ TELEFAX: (619) 535-8949
/ INFORMATION FOR SEQ ID NO: 69:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2408 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-482-279-69

Query Match          62.5%; Score 25; DB 2; Length 2408;
Best Local Similarity 100.0%; Pred.No. 0.56;
Matches 25; Conservative 0; Mismatches 0; Indels 0;

QY      16 TCATTATTGATTCAATTTGTCC 40
      |||||||
DB      495 TCATTATTGATTCAATTTGTCC 519

```

RESULT 7  
US-08-342-268-69  
; Sequence 69, Application US/08342268  
; Patent No. 5844072  
; GENERAL INFORMATION:  
; APPLICANT: Selsted, Michael E.  
; APPLICANT: Ouellette, Andre J.  
; TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods  
; TITLE OF INVENTION: of Their Use  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 18-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/930,649  
; FILING DATE: 14-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/889,020  
; FILING DATE: 26-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-UC 1206  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2408 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-342-268-69

Query Match 62.5%; Score 25; DB 2; Length 2408;  
Best Local Similarity 100.0%; Pred. No. 0.56;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCAATTGTGCC 40  
Db 495 TCATTATTGATTCAATTGTGCC 519

RESULT 8  
US-09-015-968-69  
; Sequence 69, Application US/09015968  
; Patent No. 6057425  
; GENERAL INFORMATION:  
; APPLICANT: Selsted, Michael E.  
; APPLICANT: Ouellette, Andre J.  
; TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods  
; TITLE OF INVENTION: of Their Use  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/015,968  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/482,279  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/342,268  
; FILING DATE: 18-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/930,649  
; FILING DATE: 14-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/889,020  
; FILING DATE: 26-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-UC 3003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2408 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-015-968-69

Query Match 62.5%; Score 25; DB 3; Length 2408;  
Best Local Similarity 100.0%; Pred. No. 0.56;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCAATTGTGCC 40  
Db 495 TCATTATTGATTCAATTGTGCC 519

RESULT 9  
US-09-397-386-69  
; Sequence 69, Application US/09397386  
; Patent No. 6300470  
; GENERAL INFORMATION:  
; APPLICANT: Selsted, Michael E.  
; APPLICANT: Ouellette, Andre J.  
; TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods  
; TITLE OF INVENTION: of Their Use  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/397,386  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/015,968

```
/ FILING DATE:
/ APPLICATION NUMBER: US 08/482,279
/ FILING DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/342,268
/ FILING DATE: 18-NOV-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/930,649
/ FILING DATE: 14-AUG-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/889,020
/ FILING DATE: 26-MAY-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Campbell, Cathryn A.
/ REGISTRATION NUMBER: 31,815
/ REFERENCE/DOCKET NUMBER: P-UC 3003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (619) 535-9001
/ TELEFAX: (619) 535-8949
/ INFORMATION FOR SEQ ID NO: 69:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2408 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-09-397-386-69

Query Match      62.5%; Score 25; DB 4; Length 2408;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTGTC 40
Db 495 TCATTATTGATTTCAATTGTC 519

RESULT 10
US-09-308-090-1
/ Sequence 1, Application US/09308090
/ Patent No. 6395963
/ GENERAL INFORMATION:
/ APPLICANT: Ohl, Stephan
/ APPLICANT: Van Der Lee, Frederique
/ APPLICANT: Goddijn, Oscar
/ APPLICANT: Klap, Joke
/ APPLICANT: Sijmons, Peter
/ TITLE OF INVENTION: Nematode-Inducible Regulatory DNA Sequences
/ FILE REFERENCE: MOG 57680
/ CURRENT APPLICATION NUMBER: US/09/308,090
/ CURRENT FILING DATE: 1999-05-14
/ EARLIER APPLICATION NUMBER: PCT/EP97/06472
/ EARLIER FILING DATE: 1997-11-18
/ EARLIER APPLICATION NUMBER: EP 96203213.2
/ EARLIER FILING DATE: 1996-11-18
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 3484
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (3482)..(3484)
/ US-09-308-090-1

Query Match      62.5%; Score 25; DB 4; Length 3484;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTGTC 40
Db 507 TCATTATTGATTTCAATTGTC 531

RESULT 11
US-09-380-090A-1
/ Sequence 1, Application US/09380090A
/ Patent No. 6555529
/ GENERAL INFORMATION:
/ APPLICANT: Ohl, Stephan Andreas
/ APPLICANT: Sijmons, Peter Christiaan
/ APPLICANT: Klein-Van Der Lee, Frederique
/ APPLICANT: Marianne
/ APPLICANT: Goddijn, Oscar
/ APPLICANT: Klap, Joke
/ TITLE OF INVENTION: NEMATODE-INDUCIBLE REGULATORY DNA
/ SEQUENCES
/ NUMBER OF SEQUENCES: 10
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Hale and Dorr LLP
/ STREET: 60 State Street
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: United States
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/380,090A
/ FILING DATE: 17-May-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO PCT/EP97/06472
/ FILING DATE: 18-NOV-1997
/ APPLICATION NUMBER: EP 96203213.2
/ FILING DATE: 18-NOV-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Superko, Colleen
/ REGISTRATION NUMBER: 39,850
/ REFERENCE/DOCKET NUMBER: SYN-010
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 526-6000
/ TELEFAX: (617) 526-5000
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3484 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Arabidopsis thaliana
/ STRAIN: C24
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3481..3484
/ OTHER INFORMATION: /codon_start= 3482
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1:
/ US-09-380-090A-1

Query Match      62.5%; Score 25; DB 4; Length 3484;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTGTC 40
Db 507 TCATTATTGATTTCAATTGTC 531

RESULT 12
US-09-016-366A-13
```

Sequence 13, Application US/09016366A  
Patent No. 5955431  
GENERAL INFORMATION:  
APPLICANT: Stevens, Richard L.  
APPLICANT: Huang, Chifu  
TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
TITLE OF INVENTION: INHIBITORS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02210-2211  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Fast-SEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,366A  
FILING DATE: January 30, 1998  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/037,090  
FILING DATE: 05-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7093  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
TELEX:  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3757 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-016-366A-13

Query Match 62.5%; Score 25; DB 2; Length 3757;  
Best Local Similarity 100.0%; Pred. No. 0.6;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCATTGTC 40  
Db 461 TCATTATTGATTTCATTGTC 485

RESULT 13  
US-08-978-404B-19  
Sequence 19, Application US/08978404B  
Patent No. 5968782  
GENERAL INFORMATION:  
APPLICANT: Stevens, Richard L.  
TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
TITLE OF INVENTION: FIBRINOGEN  
NUMBER OF SEQUENCES: 74  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02210-2211  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: Fast-SEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/978,404B  
FILING DATE: 25-NOV-97  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/032,354  
FILING DATE: 04-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7090  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
TELEX:  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3757 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-978-404B-19

Query Match 62.5%; Score 25; DB 2; Length 3757;  
Best Local Similarity 100.0%; Pred. No. 0.6;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCATTGTC 40  
Db 461 TCATTATTGATTTCATTGTC 485

RESULT 14  
US-08-556-978B-78/c  
Sequence 78, Application US/08556978B  
Patent No. 6268169  
GENERAL INFORMATION:  
APPLICANT: FAHNESTOCK, STEPHEN F.  
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED  
TITLE OF INVENTION: SPIDER SILK ANALOGS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSES: E. I. DU PONT DE NEMOURS AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.50 INCH  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: MICROSOFT WINDOWS 95  
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/556,978B  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/077,600  
FILING DATE: JUNE 15, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: FLOYD, LINDA AXAMETHY  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: CR-9389-A  
TELEPHONE: 302-892-8112  
TELEFAX: 302-773-0164  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4909 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
US-08-556-978B-78

Query Match 62.5%; Score 25; DB 3; Length 4909;  
Best Local Similarity 100.0%; Pred. No. 0.63;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCAATTTCGCC 40  
Db 4798 TCATTATTGATTCAATTTCGCC 4774

Search completed: December 3, 2003, 14:51:45  
Job time : 48.6033 secs

Query Match 62.5%; Score 25; DB 3; Length 4909;  
Best Local Similarity 100.0%; Pred. No. 0.63;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCAATTTCGCC 40  
Db 2288 TCATTATTGATTCAATTTCGCC 2264

RESULT 15  
US-09-068-101-7/C  
; Sequence 7, Application US/09068101  
; Patent No. 6372960  
; GENERAL INFORMATION:  
; APPLICANT: PLANT GENETIC SYSTEMS N.V.  
; TITLE OF INVENTION: Improved Barstar Gene  
; FILE REFERENCE: 2121-139P  
; CURRENT APPLICATION NUMBER: US/09/068.101  
; CURRENT FILING DATE: 1998-08-26  
; EARLIER APPLICATION NUMBER: EP 96202446.9  
; EARLIER FILING DATE: 1996-09-03  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 5349  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: other nucleic  
; OTHER INFORMATION: acid, "T-DNA of pTTS243"  
; NAME/KEY: misc feature  
; LOCATION: Complement(1)..(25)  
; OTHER INFORMATION: label = RB, "T-DNA right border"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: Complement((98)..(331))  
; OTHER INFORMATION: label = 3'g7, "region containing 3' untranslated  
; OTHER INFORMATION: end of Agrobacterium T-DNA gene 7"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: Complement((332)..(883))  
; OTHER INFORMATION: label = bar, "region coding for phosphinotricin  
; OTHER INFORMATION: acetyl transferase"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: Complement((884)..(2258))  
; OTHER INFORMATION: label = P35S, "35S promoter of Cauliflower Mosaic  
; OTHER INFORMATION: Virus"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2281)..(3969)  
; OTHER INFORMATION: label = PE1, "promoter of E1 gene of rice (WO  
; OTHER INFORMATION: 92/13956)"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (3970)..(4245)  
; OTHER INFORMATION: label = synb\*, "improved barstar DNA"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4246)..(4577)  
; OTHER INFORMATION: label = 3'chs, "region containing 3' untranslated  
; OTHER INFORMATION: end of chalcone synthase gene"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: Complement((5325)..(5349))  
; OTHER INFORMATION: label = LB, "T-DNA left border"  
US-09-068-101-7

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 13:18:51 ; Search time 178.512 Seconds  
(without alignments)  
744.732 Million cell updates/sec

Title: US-10-082-772-2

Perfect score: 40

Sequence: 1 aaggaagcgccgcctcattattgattcaattttgtcc 40

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799539 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	40	14	US-10-082-772-2
2	25	62.5	33	11	US-09-981-002-25
3	25	62.5	282	14	US-10-161-403-72
C 4	25	62.5	1763	12	US-09-244-805-57
C 5	25	62.5	1763	12	US-09-245-277-57
6	25	62.5	4346	14	US-10-161-403-113
C 7	25	62.5	5349	10	US-09-970-921-7
C 8	25	62.5	5611	10	US-09-970-921-10
9	25	62.5	38584	14	US-10-114-170-50
C 10	22.6	56.5	28313	10	US-09-764-877-3194
C 11	22.6	56.5	29228	10	US-09-764-877-3198
12	22.4	56.0	30	11	US-09-981-803-37
13	22.4	56.0	30	11	US-09-981-803-46
C 14	22.4	56.0	450	10	US-09-974-300-5700
C 15	22.2	55.5	678	12	US-10-027-632-206518
C 16	22.2	55.5	678	12	US-10-027-632-206519

C 17	22.2	55.5	678	12	US-10-027-632-206520
C 18	22.2	55.5	678	12	US-10-027-632-206521
C 19	22.2	55.5	678	13	US-10-027-632-206518
C 20	22.2	55.5	678	13	US-10-027-632-206519
C 21	22.2	55.5	678	13	US-10-027-632-206520
C 22	22.2	55.5	678	13	US-10-027-632-206521
C 23	22.2	55.5	2541	12	US-10-027-632-206522
C 24	22.2	55.5	2541	13	US-10-027-632-206522
C 25	22	55.0	34	14	US-10-161-403-4
26	22	55.0	38	9	US-09-738-363-35
C 27	22	55.0	1071	14	US-10-161-403-37
C 28	22	55.0	5855	14	US-10-161-403-112
C 29	22	55.0	5855	14	US-10-161-403-127
C 30	21.8	54.5	9827	14	US-10-114-170-66
C 31	21.4	53.5	473	12	US-09-814-353-13005
C 32	21.4	53.5	893	14	US-10-198-846-13591
33	21.2	53.0	323	10	US-09-960-352-12297
34	21	52.5	822	12	US-10-027-632-147411
35	21	52.5	822	12	US-10-027-632-147412
36	21	52.5	822	13	US-10-027-632-147411
37	21	52.5	822	13	US-10-027-632-147412
C 38	21	52.5	2441	10	US-09-822-830A-422
39	20.8	52.0	252	10	US-09-960-352-5332
C 40	20.8	52.0	558	12	US-10-027-632-191052
C 41	20.8	52.0	558	13	US-10-027-632-191052
C 42	20.8	52.0	70768	13	US-10-135-322-13
43	20.8	52.0	1691139	14	US-10-067-514-1
44	20.6	51.5	1677	10	US-09-974-300-2697
45	20.6	51.5	2732	9	US-09-822-849A-154

ALIGNMENTS

RESULT 1

US-10-082-772-2  
; Sequence 2, Application US/10082772  
; Publication No. US20030027337A1  
; GENERAL INFORMATION:  
; APPLICANT: DROGE, PETER  
; APPLICANT: CHRIST, NICOLE  
; APPLICANT: LOREACH, ELKE  
; TITLE OF INVENTION: SEQUENCE-SPECIFIC DNA RECOMBINATION IN EUKARYOTIC CELLS  
; FILE REFERENCE: DEBE:00808  
; CURRENT APPLICATION NUMBER: US/10/082,772  
; CURRENT FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: PCT/DE 00/02947  
; PRIOR FILING DATE: 2000-08-29  
; PRIOR APPLICATION NUMBER: DE 199 41 186.7  
; PRIOR FILING DATE: 1999-08-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 40  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Primer  
US-10-082-772-2

Query Match 100.0%; Score 40; DB 14; Length 40;  
Best Local Similarity 100.0%; Pred. No. 7.8e-06;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAGGAAGCGCGCTCATTATTGATTTCATTGTC 40  
Db 1 AAGGAAGCGCGCTCATTATTGATTTCATTGTC 40

RESULT 2

US-09-981-002-25  
; Sequence 25, Application US/09981002

```
; Publication No. US20030049634A1
; GENERAL INFORMATION:
; APPLICANT: Takara Shuzo Co., Ltd.
; TITLE OF INVENTION: DNA POLYMERASES WITH ENHANCED LENGTH OF PRIMER EXTENSION
; FILE REFERENCE: TR2050.1
; CURRENT APPLICATION NUMBER: US/09/981,002
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: US 08/021,623
; PRIOR FILING DATE: 1993-02-19
; PRIOR APPLICATION NUMBER: US 08/483,535
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/931,818
; PRIOR FILING DATE: 1997-09-16
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Bacteriophage lambda
US-09-981-002-25

Query Match          62.5%; Score 25; DB 11; Length 33;
Best Local Similarity 100.0%; Pred.No. 4.1;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCAATTGTGTC 40
Db 2 TCATTATTGATTCAATTGTGTC 26

RESULT 3
US-10-161-403-72
; Sequence 72, Application US/10161403
; Publication No. US20030119104A1
; GENERAL INFORMATION:
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
; FILE REFERENCE: 24601-420
; CURRENT APPLICATION NUMBER: US/10/161,403
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: atp
US-10-161-403-72

Query Match          62.5%; Score 25; DB 14; Length 282;
Best Local Similarity 100.0%; Pred.No. 6.9;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCAATTGTGTC 40
Db 243 TCATTATTGATTCAATTGTGTC 267

RESULT 4
US-09-244-805-57/c
; Sequence 57, Application US/09244805
```

```
; Publication No. US20030203840A1
; GENERAL INFORMATION:
; APPLICANT: Worley, Paul F.
; APPLICANT: Lanahan, Anthony
; APPLICANT: Goetz, Bernard
; APPLICANT: Heimisch, Holger
; APPLICANT: Kuner, Rohini
; APPLICANT: Scheek, Sigrid
; APPLICANT: Nikolich, Karoly
; APPLICANT: Zhukovski, Eugene
; TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 10496/004001
; CURRENT APPLICATION NUMBER: US/09/244,805
; CURRENT FILING DATE: 1999-02-05
; EARLIER FILING DATE: 1998-02-12
; EARLIER APPLICATION NUMBER: 60/074,518
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 1763
; TYPE: DNA
; ORGANISM: Eukaryote
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1763)
; OTHER INFORMATION: Y = C or T
; OTHER INFORMATION: n = A,T,C or G
US-09-244-805-57

Query Match          62.5%; Score 25; DB 12; Length 1763;
Best Local Similarity 100.0%; Pred.No. 11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCAATTGTGTC 40
Db 203 TCATTATTGATTCAATTGTGTC 179

RESULT 5
US-09-245-277-57/c
; Sequence 57, Application US/09245277
; Publication No. US20030211984A1
; GENERAL INFORMATION:
; APPLICANT: Worley, Paul F.
; APPLICANT: Lanahan, Anthony
; TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: JHU1530-3
; CURRENT APPLICATION NUMBER: US/09/245,277
; CURRENT FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/074,518
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: 60/074,135
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 1763
; TYPE: DNA
; ORGANISM: Eukaryote
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1763)
; OTHER INFORMATION: Y = C or T
; OTHER INFORMATION: n = A,T,C or G
US-09-245-277-57

Query Match          62.5%; Score 25; DB 12; Length 1763;
Best Local Similarity 100.0%; Pred.No. 11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```

QY      16 TCATTATTGATTTCAATTGTGCC 40
|||||
Db      203 TCATTATTGATTTCAATTGTGCC 179

```

## RESULT 6

```

US-10-161-403-113
; Sequence 113, Application US/10161403
; Publication No. US20030119104A1
; GENERAL INFORMATION:
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
; FILE REFERENCE: 2428-0120
; CURRENT APPLICATION NUMBER: US/10/161.403
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113
; LENGTH: 4346
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pSV40-193attpsensePur Plasmid
US-10-161-403-113

```

```

Query Match      62.5%; Score 25; DB 14; Length 4346;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      16 TCATTATTGATTTCAATTGTGCC 40
|||||
Db      4298 TCATTATTGATTTCAATTGTGCC 4322

```

## RESULT 7

```

US-09-970-921-7/c
; Sequence 7, Application US/09970921
; Patent No. US20020133845A1
; GENERAL INFORMATION:
; APPLICANT: Frank Michiels et al.
; TITLE OF INVENTION: Improved Barstar Gene
; FILE REFERENCE: 2428-0108P
; CURRENT APPLICATION NUMBER: US/09/970,921
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 5349
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: other nucleic
; OTHER INFORMATION: acid, "T-DNA of pRTS243"
; NAME/KEY: misc_feature
; LOCATION: Complement(11)..(25))
; OTHER INFORMATION: label = RB, "T-DNA right border"
; NAME/KEY: misc_feature
; LOCATION: Complement(198)..(331))
; OTHER INFORMATION: label = 3'g7, "region containing 3' untranslated
; OTHER INFORMATION: end of Agrobacterium T-DNA gene 7"
; NAME/KEY: misc_feature

```

```

; LOCATION: Complement((332)..(883))
; OTHER INFORMATION: label = bar, "region coding for phosphinothricin
; OTHER INFORMATION: acetyl transferase"
; NAME/KEY: misc_feature
; LOCATION: Complement((884)..(2258))
; OTHER INFORMATION: label = P35S, "35S promoter of Cauliflower Mosaic
; OTHER INFORMATION: Virus"
; NAME/KEY: misc_feature
; LOCATION: (2281)..(3989)
; OTHER INFORMATION: label = PEI, "promoter of E1 gene of rice (WO
; OTHER INFORMATION: 92/13956)"
; NAME/KEY: misc_feature
; LOCATION: (3970)..(4245)
; OTHER INFORMATION: label = synb*, "improved barstar DNA"
; NAME/KEY: misc_feature
; LOCATION: (4246)..(4577)
; OTHER INFORMATION: label = 3'chs, "region containing 3' untranslated
; OTHER INFORMATION: end of chalcone synthase gene"
; NAME/KEY: misc_feature
; LOCATION: Complement((5325)..(5349))
; OTHER INFORMATION: label = LB, "T-DNA left border"
US-09-970-921-7

```

```

Query Match      62.5%; Score 25; DB 10; Length 5349;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      16 TCATTATTGATTTCAATTGTGCC 40
|||||
Db      4798 TCATTATTGATTTCAATTGTGCC 4774

```

## RESULT 8

```

US-09-970-921-10/c
; Sequence 10, Application US/09970921
; Patent No. US20020133845A1
; GENERAL INFORMATION:
; APPLICANT: Frank Michiels et al.
; TITLE OF INVENTION: Improved Barstar Gene
; FILE REFERENCE: 2428-0108P
; CURRENT APPLICATION NUMBER: US/09/970,921
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 5611
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: other nucleic
; OTHER INFORMATION: acid, "plasmid pLH48"
; NAME/KEY: misc_feature
; LOCATION: Complement((39)..(317))
; OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
; OTHER INFORMATION: end of nopaline synthase gene of Agrobacterium
; OTHER INFORMATION: T-DNA"
; NAME/KEY: misc_feature
; LOCATION: Complement((318)..(869))
; OTHER INFORMATION: label = bar, "region coding for phosphinothricin
; OTHER INFORMATION: acetyl transferase"
; NAME/KEY: misc_feature
; LOCATION: Complement((870)..(1702))
; OTHER INFORMATION: label = P35S, "35S promoter of Cauliflower Mosaic
; OTHER INFORMATION: Virus"
; NAME/KEY: misc_feature
; LOCATION: (1740)..(2284)
; OTHER INFORMATION: label = PTA29, "promoter of TA29 gene of Nicotiana
; OTHER INFORMATION: tabacum"
; NAME/KEY: misc_feature
; LOCATION: (2285)..(2560)
; OTHER INFORMATION: label = synb*, "improved barstar DNA"
; NAME/KEY: misc_feature
; LOCATION: (2561)..(2892)

```

```
; OTHER INFORMATION: label = 3'chs, "region containing 3' untranslated
; OTHER INFORMATION: end of chalcone synthase gene"
US-09-970-921-10

Query Match      62.5%; Score 25; DB 10; Length 5611;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCAATTGGCC 40
Db 3113 TCATTATTGATTCAATTGGCC 3089

RESULT 9
US-10-114-170-50
; Sequence 50, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. US20030023075A1 Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,170
; FILING DATE: 01-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/453,702
; FILING DATE: 03-DEC-1999
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38584
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 50:

US-10-114-170-50

Query Match      62.5%; Score 25; DB 14; Length 38584;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCAATTGGCC 40
Db 82 TCATTATTGATTCAATTGGCC 106

RESULT 10
```

```
US-09-764-877-3194/C
; Sequence 3194, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3194
; LENGTH: 28313
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3194

Query Match      56.5%; Score 22.6; DB 10; Length 28313;
Best Local Similarity 75.7%; Pred. No. 1.7e+02;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 AAGGAAGCGCGCTCATTATTGATTTCATTTG 37
Db 16691 AAGGAAGCGCGCTCATTACCTGATTACCAATTTG 16655

RESULT 11
US-09-764-877-3198/C
; Sequence 3198, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3198
; LENGTH: 29228
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3198

Query Match      56.5%; Score 22.6; DB 10; Length 29228;
Best Local Similarity 75.7%; Pred. No. 1.7e+02;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 AAGGAAGCGCGCTCATTATTGATTTCATTTG 37
Db 14237 AAGGAAGCGCGCTCATTACCTGATTACCAATTTG 14201

RESULT 12
US-09-981-803-37
; Sequence 37, Application US/09981803
; Publication No. US20030032092A1
; GENERAL INFORMATION:
; APPLICANT: Joel CROUZET
; APPLICANT: Daniel SCHERMAN
; APPLICANT: Beatrice CAMERON
; APPLICANT: Pierre WILLS
; APPLICANT: Anne-Marie DARQUET
; TITLE OF INVENTION: DNA MOLECULES, PREPARATION AND USE IN GENE THERAPY
; FILE REFERENCE: MINICIRCLE
; CURRENT APPLICATION NUMBER: US/09/981,803
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 30
```

```

; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence:
; OTHER INFORMATION: oligonucleotide
US-09-981-803-37

```

Query Match 56.0%; Score 22.4; DB 11; Length 30;  
Best Local Similarity 95.8%; Pred. No. 40;  
Matches 23; Conservative 0; Mismatches 1; Indels

**Qy** 14 GCTCATTATTTGATTTCAAATTTG 37  
| | | | |  
**Dd** 7 GGTCATTATTTGATTTCAAATTTG 30

```

RESULT 13
US-09-981-803-46
; Sequence 46, Application US/09981803
; Publication No. US20030032092A1
; GENERAL INFORMATION:
; APPLICANT: JOEL CROUZET
; APPLICANT: DANIEL SCHERMAN
; APPLICANT: BEATRICE CAMERON
; APPLICANT: PIERRE WILS
; APPLICANT: ANNE-MARIE DARQUET
; TITLE OF INVENTION: DNA MOLECULES, PREPARATION AND USE IN GENE THERAPY
; FILE REFERENCE: MINICIRCLE
; CURRENT APPLICATION NUMBER: US/09/981,803
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence:
; - OTHER INFORMATION: oligonucleotide
US-09-981-803-46

```

Query Match 56.0%; Score 22.4; DB 11; Length 30;  
Best Local Similarity 95.8%; Pred. NO. 40;  
Matches 23; Conservative 0; Mismatches 1; Indels

**QY** 14 GCTCATTATTTGATTTCAAATTTTG 37  
| | | | |  
**Dd** 7 GGTCATTATTTGATTTCAAATTTTG 30

```

RESULT 14
US-09-974-300-5700/c
; Sequence 5700, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 5700
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-5700

```

Query Match	56.0%	Score 22.4;	DB 10;	Length 450;
Best Local Similarity	72.5%	Pred. No. 76;		
Matches 29; Conservative	0;	Mismatches 11;	Indels 0	

[illegible]

```

RESULT 15
US-10-027-632-206518/c
; Sequence 206518, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 206518
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(678)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-206518

```

Query Match	55.5%	Score 22.2;	DB 12;	Length 678;
Best Local Similarity	77.1%	Pred. No. 1e+02;		
Matches 27; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;

**Qy** 1 AAGGAAAGCGGCCGTCAATTATTTGATTTCAAATT 35  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

**Db** 295 AAGAATAGTCCCTGCACATTTTGATTTGATTCGATTT 261

Search completed: December 3, 2003, 17:22:24  
Job time : 180.512 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 1793.72 Seconds  
(without alignments)  
541.991 Million cell updates/sec

Title: US-10-082-772-2

Perfect score: 40

Sequence: 1 aagaaagcgccgcctcattattgatttcattttgtcc 40

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
C	1	25	62.5	206	13	BQ156416
	2	25	62.5	360	13	AU244794
C	3	25	62.5	388	13	BQ156182
	4	25	62.5	395	28	AQ991303

5	25	62.5	417	12	BM134283
6	25	62.5	472	13	BQ157398
7	25	62.5	473	13	BQ156404
8	25	62.5	542	10	BG12744
9	25	62.5	572	28	AQ989502
10	25	62.5	587	28	AQ991064
11	25	62.5	597	12	B1422679
12	25	62.5	622	10	B435088
13	25	62.5	641	28	AQ990513
14	25	62.5	664	28	AQ991011
15	25	62.5	672	28	AQ990864
16	25	62.5	675	28	AQ991241
17	25	62.5	695	28	AQ991039
18	25	62.5	712	28	AQ990809
19	25	62.5	719	28	AQ991352
20	25	62.5	743	28	AQ990346
21	25	62.5	746	28	AQ990173
22	25	62.5	751	28	AQ989566
23	25	62.5	753	28	AQ990861
24	25	62.5	764	28	AQ990110
25	25	62.5	767	28	AQ990878
26	25	62.5	767	28	AQ990301
27	25	62.5	769	28	AQ990470
28	25	62.5	808	28	AQ990388
29	24.8	62.0	517	12	BM533719
30	24.4	61.0	958	29	CNS04HMU
31	24.2	60.5	370	28	AQ193727
32	24.2	60.5	907	10	BF106073
33	24	60.0	854	28	AQ740016
34	23.4	58.5	851	28	AQ990758
35	23	57.5	698	28	BH730850
36	23	57.5	811	28	BH544112
37	22.8	57.0	481	9	AA252744
38	22.8	57.0	750	14	CB292689
39	22.6	56.5	631	29	CNS01NCR
40	22.6	56.5	704	29	BZ773242
41	22.6	56.5	839	14	CB565198
42	22.4	56.0	461	28	AQ664509
43	22.4	56.0	657	10	B5624763
44	22.4	56.0	1095	12	BM804116
45	22.2	55.5	499	14	CB092932

#### ALIGNMENTS

RESULT 1  
BQ156416/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Medicago truncatula (barrel medic)

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

euromids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE

AUTHORS

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula irradiated library

Unpublished

JOURNAL

COMMENT

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7380

Email: gdmay@noble.org

BQ156416 206 bp mRNA linear EST 24-APR-2002  
NF092F02IR1F1027 Irradiated Medicago truncatula cDNA clone  
NF092F02IR 5', mRNA sequence.

BQ156416  
NF092F02IR 5', mRNA sequence.  
BQ156416  
EST.  
Medicago truncatula (barrel medic)  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
euromids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

1 (bases 1 to 206)  
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula irradiated library  
Unpublished  
Contact: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org

BQ156416 206 bp mRNA linear EST 24-APR-2002  
NF092F02IR 5', mRNA sequence.  
BQ156416  
EST.  
Medicago truncatula (barrel medic)  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
euromids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

```

Insert Length: 206 Std Error: 0.00
Plate: 092 row: F column: 02
Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES
    source
        1..206
            /organism="Medicago truncatula"
            /mol_type="mRNA"
            /db_xref="taxon:3880"
            /clone="NF092F02IR"
            /tissue_type="seedlings"
            /dev_stage="seedling"
            /clone_lib="irradiated"
            /note="Vector: Lambda Zap; Seedlings were exposed either
            to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
            Gamma-irradiated samples were harvested at 6, 12, 24 and
            48 hours after treatment. UV-irradiated samples were
            harvested 24 hours post-treatment. cDNA was prepared from
            polyA+ enriched, pooled samples of equivalent amounts of
            total RNA from each sample. The cDNA was directionally
            ligated into the Uni-Zap XR vector (Stratagene) and
            packaged using the Gigapack III Gold packaging extracts.
            Phagemids containing cDNA inserts were in vivo excised
            from the recombinant Uni-Zap XR vector using ExAssist
            helper phage and the E. coli strain XL1-Blue MRF'
            (Stratagene). Excised plasmids were plated using SOLR
            cells."

BASE COUNT      81 a  27 c  39 g  59 t
ORIGIN
Query Match      62.5%; Score 25; DB 13; Length 206;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTTCATTGTC 40
Db 77 TCATTATTGATTTCATTGTC 53

RESULT 2
LOCUS      AU244794      360 bp      mRNA      linear      EST 21-FEB-2002
DEFINITION AU244794 Shibata Xenopus AEM lambda-ZAP II cDNA library Xenopus
            laevis cDNA clone p6e6 5', mRNA sequence.
ACCESSION  AU244794
VERSION     AU244794.1 GI:18850722
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
ORGANISM    Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1 (bases 1 to 360)
            Shibata,M., Itoh,M., Ohmori,S., Shinga,J. and Taira,M.
            Systematic screening and expression analysis of the head organizer
            genes in Xenopus embryos
            Dev. Biol. 239 (2), 241-256 (2001)
            PubMed 11784032
COMMENT     Contact: Masanori Taira
            Department of Biological Sciences
            Graduate School of Science, University of Tokyo; CREST, Japan
            Science and Technology Corporation, Japan
            7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
            Tel: 81-03-5841-4434
            Fax: 81-03-5841-4434
            Email: m.taira@biol.s.u-tokyo.ac.jp,
            URL: http://www.biol.s.u-tokyo.ac.jp/users/lmb/lmb-hp.html.
            Location/Qualifiers
                1..360
                    /organism="Xenopus laevis"
                    /mol_type="mRNA"
                    /db_xref="taxon:8355"
                    /clone="p6e6"

FEATURES
    source
        1..388
            /organism="Medicago truncatula"
            /mol_type="mRNA"
            /db_xref="taxon:3880"
            /clone="NF090A02IR"
            /tissue_type="seedlings"
            /dev_stage="seedling"
            /clone_lib="irradiated"
            /note="Vector: Lambda Zap; Seedlings were exposed either
            to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
            Gamma-irradiated samples were harvested at 6, 12, 24 and
            48 hours after treatment. UV-irradiated samples were
            harvested 24 hours post-treatment. cDNA was prepared from
            polyA+ enriched, pooled samples of equivalent amounts of
            total RNA from each sample. The cDNA was directionally
            ligated into the Uni-Zap XR vector (Stratagene) and
            packaged using the Gigapack III Gold packaging extracts.
            Phagemids containing cDNA inserts were in vivo excised
            from the recombinant Uni-Zap XR vector using ExAssist
            helper phage and the E. coli strain XL1-Blue MRF'
            (Stratagene). Excised plasmids were plated using SOLR
            cells."

BASE COUNT      101 a  78 c  78 g  101 t  2 others
ORIGIN
Query Match      62.5%; Score 25; DB 9; Length 360;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTTCATTGTC 40
Db 231 TCATTATTGATTTCATTGTC 255

RESULT 3
LOCUS      BQ156182      388 bp      mRNA      linear      EST 24-APR-2002
DEFINITION NF090A02IRF1017 Irradiated Medicago truncatula cDNA clone
            NF090A02IR 5', mRNA sequence.
ACCESSION  BQ156182
VERSION     BQ156182.1 GI:20293229
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
ORGANISM    Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
            1 (bases 1 to 388)
REFERENCE   1 (bases 1 to 388)
            Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
            Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula irradiated library
            Unpublished
            Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7391
            Fax: 580 221 7380
            Email: gdmay@noble.org
            Insert Length: 388 Std Error: 0.00
            Plate: 090 row: A column: 02
            Seq primer: TCACACAGGAACAGCTATGAC.
            Location/Qualifiers
                1..388
                    /organism="Medicago truncatula"
                    /mol_type="mRNA"
                    /db_xref="taxon:3880"
                    /clone="NF090A02IR"
                    /tissue_type="seedlings"
                    /dev_stage="seedling"
                    /clone_lib="irradiated"
                    /note="Vector: Lambda Zap; Seedlings were exposed either
                    to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
                    Gamma-irradiated samples were harvested at 6, 12, 24 and
                    48 hours after treatment. UV-irradiated samples were
                    harvested 24 hours post-treatment. cDNA was prepared from
                    polyA+ enriched, pooled samples of equivalent amounts of
                    total RNA from each sample. The cDNA was directionally
                    ligated into the Uni-Zap XR vector (Stratagene) and
                    packaged using the Gigapack III Gold packaging extracts.
                    Phagemids containing cDNA inserts were in vivo excised
                    from the recombinant Uni-Zap XR vector using ExAssist
                    helper phage and the E. coli strain XL1-Blue MRF'
                    (Stratagene). Excised plasmids were plated using SOLR
                    cells."

BASE COUNT      101 a  74 c  99 g  114 t
ORIGIN
Query Match      62.5%; Score 25; DB 13; Length 388;

```

Best Local Similarity 100.0%; Pred. No. 6.1e+02; Mismatches 25; Conservative 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTTCAATTTTGTC 40  
Db 171 TCATTATTGATTTCAATTTTGTC 147

RESULT 4  
AQ991303/c  
LOCUS  
DEFINITION Rfc02205 Photorhabdus luminescens strain W14 M13 library  
ACCESSION AQ991303  
VERSION AQ991303.1 GI:9649897  
KEYWORDS GSS  
SOURCE Photorhabdus luminescens  
ORGANISM Photorhabdus luminescens  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus.  
REFERENCE 1 (bases 1 to 395)  
AUTHORS french-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T., Daborn,P.J., Bowen,D. and Blattner,P.R.  
TITLE A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence  
JOURNAL Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)  
MEDLINE 20378633  
PubMed 10919786  
COMMENT Contact: french-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK  
Tel: (44) 1225 826821  
Fax: (44) 1225 826779  
Email: bsr1c@bath.ac.uk  
This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli X12 genome) please see french-Constant et al. 2000, Nucleic Acids Res.  
Seq primer: M13 Forward  
Class: shotgun.  
Location/Qualifiers  
FEATURES  
source 1..395  
/organism="Photorhabdus luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG02205"  
/dev\_stage="primary phase variant"  
/clone\_lib="Photorhabdus luminescens strain W14 M13 library"  
/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."

BASE COUNT 135 a 72 c 63 g 101 t 24 others

Query Match 62.5%; Score 25; DB 28; Length 395;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTTCAATTTTGTC 40  
Db 120 TCATTATTGATTTCAATTTTGTC 96

RESULT 5  
BM134283  
LOCUS  
DEFINITION WHE0488 B03 C062S Wheat Fusarium graminearum infected spike cDNA library,Triticum aestivum cDNA clone WHE0488\_B03\_C06, mRNA sequence.  
ACCESSION BM134283

BM134283.1 GI:17141863  
EST.  
Triticum aestivum (bread wheat)  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.  
1 (bases 1 to 417)  
Anderson,O.D., Chao,S., Han,P.S., Heinen,S., Heia,C.C., Kang,Y., Kruger,W.M., Lazo,G.R., Miller,S., Muehlbauer,G.J., Miller,R., Pritsch,C., Rausch,C.J., Seaton,C.L., Tong,J.C., Vance,C. and Wilson,C.F.  
The structure and function of the expressed portion of the wheat genomes - Fusarium graminearum infected spike cDNA library  
Unpublished  
Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105595773  
Fax: 5105595818  
Email: oanderson@pw.usda.gov  
Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20. No effort was taken to identify ESTs of fungal origin from this library, thus this EST could be of wheat or fungal origin.  
Seq primer: StrataGene SK primer.  
Location/Qualifiers  
FEATURES  
source 1..417  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="Sumai3"  
/db\_xref="taxon:4565"  
/clone="WHE0488\_B03\_C06"  
/tissue\_type="Spike"  
/dev\_stage="Adult plant"  
/lab\_host="E. coli SOLR"  
/clone\_lib="Wheat Fusarium graminearum infected spike cDNA library"  
/note="Vector: Lambda Uni-ZAP XR, excised phagemid pBluescript SK; Site 1: EcoRI; Site 2: XhoI; Plants were grown in the greenhouse. Spikes were sprayed at anthesis with Fusarium graminearum. Total RNA, and poly(A) RNA were prepared and pooled from infected spike at 0, 6, 12, 24, 36 and 48 hours after inoculation, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in G. Muehlbauer lab at the University of Minnesota (Kruger, W.M., Muehlbauer, G.J., Pritsch, C., Vance, C.). The cDNA library should contain genes of both wheat and fungal pathogen origin. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 116 a 94 c 94 g 113 t

Query Match 62.5%; Score 25; DB 12; Length 417;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTTCAATTTTGTC 40  
Db 246 TCATTATTGATTTCAATTTTGTC 270

RESULT 6  
BQ157398/c  
LOCUS  
DEFINITION NF104D07IR1F062 Irradiated Medicago truncatula cDNA clone  
ACCESSION BQ157398  
VERSION BQ157398.1 GI:20294457  
KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)

ORGANISM  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eutrosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.

REFERENCE  
1 (bases 1 to 472)  
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula irradiated library

TITLE  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Unpublished

JOURNAL  
COMMENT  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org  
Insert Length: 472 Std Error: 0.00  
Plate: 104 row: D column: 07  
Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES  
Location/Qualifiers  
1..472  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/db\_xref="taxon:3880"  
/clone="NF104D071R"  
/tissue\_type="seedlings"  
/dev\_stage="seedling"  
/clone\_lib="irradiated"  
/notes="Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

BASE COUNT  
162 a 89 c 95 g 126 t

ORIGIN  
Query Match 62.5%; Score 25; DB 13; Length 472;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCGAATTTTGCC 40  
|||||  
Db 290 TCATTATTGATTCGAATTTTGCC 266  
|||||

RESULT 7  
BQ156404/c  
LOCUS  
DEFINITION  
NF092E031R1F1023 Irradiated Medicago truncatula cDNA clone  
BQ156404  
BQ156404.1 GI:20293463  
EST.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Medicago truncatula (barrel medic)  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eutrosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.

REFERENCE  
1 (bases 1 to 473)  
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

TITLE  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula irradiated library  
Unpublished

JOURNAL  
COMMENT  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org  
Insert Length: 473 Std Error: 0.00  
Plate: 092 row: E column: 03  
Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES  
Location/Qualifiers  
1..473  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/db\_xref="taxon:3880"  
/clone="NF092E031R"  
/tissue\_type="seedlings"  
/dev\_stage="seedling"  
/clone\_lib="irradiated"  
/notes="Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

BASE COUNT  
162 a 90 c 95 g 126 t

ORIGIN  
Query Match 62.5%; Score 25; DB 13; Length 473;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCGAATTTTGCC 40  
|||||  
Db 291 TCATTATTGATTCGAATTTTGCC 267  
|||||

RESULT 8  
BG312744  
LOCUS  
DEFINITION  
WHE2457 F05 L092S Triticum monococcum cDNA clone WHE2457\_F05\_L09, mRNA  
BG312744  
BG312744.1 GI:13114547  
EST.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Triticum monococcum  
Triticum monococcum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae ; Triticeae; Triticum.  
1 (bases 1 to 542)  
Anderson, O.D., Chao, S., Dubcovsky, J., Echenique, V., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L., Stamova, B. and Tong, J.C.  
The structure and function of the expressed portion of the wheat genomes - Early reproductive apex cDNA library from Triticum monococcum  
Unpublished

REFERENCE  
1 (bases 1 to 542)  
Anderson, O.D., Chao, S., Dubcovsky, J., Echenique, V., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L., Stamova, B. and Tong, J.C.  
The structure and function of the expressed portion of the wheat genomes - Early reproductive apex cDNA library from Triticum monococcum  
Unpublished

JOURNAL  
COMMENT  
Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105595773  
Fax: 5105595818  
Email: oanderson@pw.usda.gov  
Sequence have been trimmed to remove vector sequence and low  
quality sequence with phred score less than 20  
Seq primer: Stratagene SK primer.

FEATURES

source

Location/Qualifiers  
1..542  
/organism="Triticum monococcum"  
/mol\_type="mRNA"  
/cultiivar="DV92"  
/db\_xref="taxon:4568"  
/clone="WHE2457\_F05\_L09"  
/tissue type="Early reproductive apex"  
/dev stage="Seven week-old plants"  
/lab\_host="E. coli XL0LR"  
/clone\_lib="Triticum monococcum early reproductive apex  
cDNA library"  
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
Site 1: EcoRI; Site 2: XhoI; The tissue, total RNA, and  
poly(A) RNA were prepared from apex at double-ridge stage  
to terminal-spikelet stage during transition from  
vegetative state to flower state, a cDNA library was made,  
and the cDNA clones were in vivo excised at the  
University of California, Davis (V. Echenique, B. Stamova  
, J. Dubcovsky). Plasmid DNA preparations and DNA  
sequencing were performed in the OD Anderson lab (all  
other authors)."

BASE COUNT 141 a 129 c 107 g 164 t 1 others  
ORIGIN  
Query Match 62.5%; Score 25; DB 10; Length 542;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTTGTCC 40  
Db 35 TCATTATTGATTTCAATTTGTCC 59

RESULT 9

AQ989502/c

LOCUS

DEFINITION

Photobacterium luminescens strain W14 M13 library

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Photobacterium luminescens

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Photobacterium.

1 (bases 1 to 572)

ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,

Daborn, P.J., Bowen, D. and Blattner, F.R.

A genomic sample sequence of the entomopathogenic bacterium

Photobacterium luminescens W14: potential implications for virulence

Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

Seq primer: M13 Forward  
Class: shotgun.

FEATURES

source

Location/Qualifiers  
1..572  
/organism="Photobacterium luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG00051"  
/dev stage="primary phase variant"  
/clone\_lib="Photobacterium luminescens strain W14 M13  
library"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."

BASE COUNT 143 a 130 c 144 g 153 t 2 others  
ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 572;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTTGTCC 40  
Db 534 TCATTATTGATTTCAATTTGTCC 510

RESULT 10

AQ991064/c

LOCUS

DEFINITION

Photobacterium luminescens strain W14 M13 library

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Photobacterium luminescens

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Photobacterium.

1 (bases 1 to 587)

ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,

Daborn, P.J., Bowen, D. and Blattner, F.R.

A genomic sample sequence of the entomopathogenic bacterium

Photobacterium luminescens W14: potential implications for virulence

Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

20378633

PUBLISHED

COMMENT

Contact: ffrench-Constant RH

Department of Biology and Biochemistry

University of Bath

South Building, Bath BA2 7AY, UK

Seq primer: M13 Forward  
Class: shotgun.

FEATURES

source

Location/Qualifiers  
1..587  
/organism="Photobacterium luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG01922"  
/dev stage="primary phase variant"  
/clone\_lib="Photobacterium luminescens strain W14 M13  
library"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."

BASE COUNT 150 a 132 c 147 g 156 t 2 others  
ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 572;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTTGTCC 40  
Db 534 TCATTATTGATTTCAATTTGTCC 510

RESULT 10

AQ991064/c

LOCUS

DEFINITION

Photobacterium luminescens strain W14 M13 library

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Photobacterium luminescens

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Photobacterium.

1 (bases 1 to 587)

ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,

Daborn, P.J., Bowen, D. and Blattner, F.R.

A genomic sample sequence of the entomopathogenic bacterium

Photobacterium luminescens W14: potential implications for virulence

Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

20378633

PUBLISHED

COMMENT

Contact: ffrench-Constant RH

Department of Biology and Biochemistry

University of Bath

South Building, Bath BA2 7AY, UK

Seq primer: M13 Forward  
Class: shotgun.

FEATURES

source

Location/Qualifiers  
1..572  
/organism="Photobacterium luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG00051"  
/dev stage="primary phase variant"  
/clone\_lib="Photobacterium luminescens strain W14 M13  
library"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."

BASE COUNT 143 a 130 c 144 g 153 t 2 others  
ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 572;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTTGTCC 40  
Db 534 TCATTATTGATTTCAATTTGTCC 510

RESULT 10

AQ991064/c

LOCUS

DEFINITION

Photobacterium luminescens strain W14 M13 library

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Photobacterium luminescens

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Photobacterium.

1 (bases 1 to 587)

ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,

Daborn, P.J., Bowen, D. and Blattner, F.R.

A genomic sample sequence of the entomopathogenic bacterium

Photobacterium luminescens W14: potential implications for virulence

Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

20378633

PUBLISHED

COMMENT

Contact: ffrench-Constant RH

Department of Biology and Biochemistry

University of Bath

South Building, Bath BA2 7AY, UK

Seq primer: M13 Forward  
Class: shotgun.

FEATURES

source

Location/Qualifiers  
1..572  
/organism="Photobacterium luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG00051"  
/dev stage="primary phase variant"  
/clone\_lib="Photobacterium luminescens strain W14 M13  
library"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."

BASE COUNT 143 a 130 c 144 g 153 t 2 others  
ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 572;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTTGTCC 40  
Db 534 TCATTATTGATTTCAATTTGTCC 510

RESULT 10

AQ991064/c

LOCUS

DEFINITION

Photobacterium luminescens strain W14 M13 library

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Photobacterium luminescens

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Photobacterium.

1 (bases 1 to 587)

ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,

Daborn, P.J., Bowen, D. and Blattner, F.R.

A genomic sample sequence of the entomopathogenic bacterium

Photobacterium luminescens W14: potential implications for virulence

Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

20378633

PUBLISHED

COMMENT

Contact: ffrench-Constant RH

Department of Biology and Biochemistry

University of Bath

South Building, Bath BA2 7AY, UK

Seq primer: M13 Forward  
Class: shotgun.

FEATURES

source

Location/Qualifiers  
1..572  
/organism="Photobacterium luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG00051"  
/dev stage="primary phase variant"  
/clone\_lib="Photobacterium luminescens strain W14 M13  
library"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."

BASE COUNT 143 a 130 c 144 g 153 t 2 others  
ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 572;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTTGTCC 40  
Db 534 TCATTATTGATTTCAATTTGTCC 510

RESULT 10

AQ991064/c

LOCUS

DEFINITION

Photobacterium luminescens strain W14 M13 library

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Photobacterium luminescens

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Photobacterium.

1 (bases 1 to 587)

ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,

Daborn, P.J., Bowen, D. and Blattner, F.R.

A genomic sample sequence of the entomopathogenic bacterium

Photobacterium luminescens W14: potential implications for virulence

Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

20378633

PUBLISHED

COMMENT

Contact: ffrench-Constant RH

Department of Biology and Biochemistry

University of Bath

South Building, Bath BA2 7AY, UK

Seq primer: M13 Forward  
Class: shotgun.

FEATURES

source

Location/Qualifiers  
1..572  
/organism="Photobacterium luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG00051"  
/dev stage="primary phase variant"  
/clone\_lib="Photobacterium luminescens strain W14 M13  
library"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."

BASE COUNT 143 a 130 c 144 g 153 t 2 others  
ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 572;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTTGTCC 40  
Db 534 TCATTATTGATTTCAATTTGTCC 510

RESULT 10

AQ991064/c

LOCUS

DEFINITION

Photobacterium luminescens strain W14 M13 library

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Photobacterium luminescens

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Photobacterium.

1 (bases 1 to 587)

ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,

Daborn, P.J., Bowen, D. and Blattner, F.R.

A genomic sample sequence of the entomopathogenic bacterium

Photobacterium luminescens W14: potential implications for virulence

Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

20378633

PUBLISHED

COMMENT

Contact: ffrench-Constant RH

Department of Biology and Biochemistry

University of Bath

South Building, Bath BA2 7AY, UK

Seq primer: M13 Forward  
Class: shotgun.

FEATURES

source

Location/Qualifiers  
1..572  
/organism="Photobacterium luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG00051"  
/dev stage="primary phase variant"  
/clone\_lib="Photobacterium luminescens strain W14 M13  
library"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."

BASE COUNT 143 a 130 c 144 g 153 t 2 others  
ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 572;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCAATTTGTCC 40  
Db 534 TCATTATTGATTTCAATTTGTCC 510

RESULT 10

AQ991064/c

LOCUS

DEFINITION

Photobacterium luminescens strain W14 M13 library

sequence.

ACCESSION



## ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 587;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-02;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCATTTGTC 40  
 |||||  
 Db 517 TCATTATTGATTCATTTGTC 493

## RESULT 11

BI422679/c 597 bp mRNA linear EST 16-AUG-2001  
 LOCUS EST533345 tomato callus, TAMU Lycopersicon esculentum cDNA clone  
 DEFINITION cLEC71G2 5' end, mRNA sequence.

ACCESSION BI422679  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 597)  
 AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,  
 Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning  
 , C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.

TITLE Generation of ESTs from tomato callus tissue

JOURNAL Unpublished

COMMENT Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>.

## FEATURES

source Location/Qualifiers

1..597  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLEC71G2"  
 /tissue\_type="callus"  
 /dev\_stage="25-40 days old"  
 /lab\_host="XLI-Blue MRF"  
 /clone\_lib="tomato callus, TAMU"  
 /notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; supplier: Giovannoni laboratory; CLSC - Cotyledons  
 of seedlings 7-10 days post-germination were excised, cut  
 at both ends and placed on MS medium with no selection.  
 Mixed callus was harvested at 25 and 40 days and included  
 undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 193 a 109 c 131 g 164 t

## ORIGIN

Query Match 62.5%; Score 25; DB 12; Length 597;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-02;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCATTTGTC 40  
 |||||  
 Db 462 TCATTATTGATTCATTTGTC 438

## RESULT 12

BE435088 622 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST406166 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
 DEFINITION clone cLEG25A5, mRNA sequence.

ACCESSION BE435088  
 VERSION BE435088.1 GI:9432931  
 KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

## ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

## REFERENCE

## AUTHORS

Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
 Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,  
 Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley

,S.D.

Generation of ESTs from tomato fruit tissue, breaker stage

Unpublished

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

## FEATURES

source Location/Qualifiers

1..622  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLEG25A5"  
 /tissue\_type="Pericarp"  
 /dev\_stage="breaker"  
 /lab\_host="SOLR"  
 /clone\_lib="tomato breaker fruit, TIGR"  
 /notes="Vector: pBluescriptSKCudapt; Site 1: EcoRI;  
 Site 2: XhoI; Fruit were harvested at the breaker stage  
 (first sign of lycopene accumulation on the blossom end of  
 the fruit). Fruit were cut in half and the seeds and  
 locules were discarded prior to freezing the pericarp."

BASE COUNT 168 a 143 c 145 g 166 t

## ORIGIN

Query Match 62.5%; Score 25; DB 10; Length 622;

Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 TCATTATTGATTCATTTGTC 40  
 |||||  
 Db 261 TCATTATTGATTCATTTGTC 285

## RESULT 13

## LOCUS

AQ990513/c 641 bp DNA linear GSS 14-AUG-2000  
 DEFINITION Rfcl1299 Photorhabdus luminescens strain W14 M13 library  
 Photorhabdus luminescens genomic clone PLG01299, genomic survey  
 sequence.

ACCESSION AQ990513

VERSION AQ990513.1 GI:9649107

## KEYWORDS

## SOURCE

## ORGANISM

Photorhabdus luminescens  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Photorhabdus.

## REFERENCE

## AUTHORS

1 (bases 1 to 641)  
 fFrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,  
 Daborn,P.J., Bowen,D. and Blattner,F.R.

A genomic sample sequence of the entomopathogenic bacterium

Photorhabdus luminescens W14: potential implications for virulence

Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

20378633

JOURNAL 10919786

## MEDLINE

## PUBMED

## COMMENT

Contact: fFrench-Constant RH

Department of Biology and Biochemistry

University of Bath

South Building, Bath BA2 7AY, UK

Tel: (44) 1225 826621

Fax: (44) 1225 826779

Email: [bsrffc@bath.ac.uk](mailto:bsrffc@bath.ac.uk)

This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res.

Seq primer: M13 Forward

Class: shotgun.

#### FEATURES

source  
Location/Qualifiers  
1..641  
/organism="Photorhabdus luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG01239"  
/dev\_stage="primary phase variant"  
/clone\_lib="Photorhabdus luminescens strain W14 M13 library"  
/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."

BASE COUNT 167 a 145 c 158 g 170 t 1 others

ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 641;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCATTGTC 40  
Db 545 TCATTATTGATTTCATTGTC 521

RESULT 14  
AQ991011/c  
LOCUS  
DEFINITION  
Photorhabdus luminescens strain W14 M13 library  
sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Contact: french-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK  
Tel: (44) 1225 826621  
Fax: (44) 1225 826779  
Email: bssr@bath.ac.uk  
This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res.

Seq primer: M13 Forward

Class: shotgun.

#### FEATURES

source  
Location/Qualifiers  
1..864  
/organism="Photorhabdus luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG01864"  
/dev\_stage="primary phase variant"  
/clone\_lib="Photorhabdus luminescens strain W14 M13 library"

#### library"

/notes="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."

BASE COUNT 178 a 146 c 162 g 175 t 3 others

ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 664;  
Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCATTGTC 40  
Db 548 TCATTATTGATTTCATTGTC 524

RESULT 15  
AQ990864/c  
LOCUS  
DEFINITION

AQ990864 672 bp DNA linear GSS 14-AUG-2000  
Rfco1701 Photorhabdus luminescens strain W14 M13 library  
Photorhabdus luminescens genomic clone PLG01701, genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Contact: french-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK  
Tel: (44) 1225 826621  
Fax: (44) 1225 826779  
Email: bssr@bath.ac.uk  
This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res.

Seq primer: M13 Forward

Class: shotgun.

#### FEATURES

source  
Location/Qualifiers  
1..672  
/organism="Photorhabdus luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG01701"  
/dev\_stage="primary phase variant"  
/clone\_lib="Photorhabdus luminescens strain W14 M13 library"  
/notes="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."

BASE COUNT 184 a 146 c 161 g 179 t 2 others

ORIGIN

Query Match 62.5%; Score 25; DB 28; Length 672;  
Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TCATTATTGATTTCATTGTC 40  
Db 549 TCATTATTGATTTCATTGTC 525

Search completed: December 3, 2003, 14:49:16  
Job time : 1796.72 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 652.562 Seconds  
(without alignments)  
1504.579 Million cell updates/sec

Title: US-10-082-772-3

Perfect score: 24

Sequence: 1 gttcagcttttgataactaagtg 24

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2889711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_nam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	24	100.0	243	6	AX092116	AX092116 Sequence
2	22.4	93.3	201	6	AR044609	AR044609 Sequence
3	22.4	93.3	201	6	E05439	E05439 Oligonucleo
4	22.4	93.3	201	6	I13139	I13139 Sequence 18
5	22.4	93.3	201	6	I36498	I36498 Sequence 18
6	22.4	93.3	243	6	AX092113	AX092113 Sequence
7	22.4	93.3	361	7	LAMINTATT	M23841 Bacterioph
8	22.4	93.3	610	6	AX101000	AX101000 Sequence
9	22.4	93.3	1668	9	MACHSS	D85521 Macaca fasc
10	22.4	93.3	2758	3	PFAHRPC	M17028 P.falciapar
11	22.4	93.3	2959	12	AY048723	AY048723 CRIM plas
12	22.4	93.3	3485	12	AF178449	AF178449 Integrati
13	22.4	93.3	3485	12	AF178450	AF178450 Integrati
14	22.4	93.3	3663	12	AY048716	AY048716 CRIM plas
15	22.4	93.3	3695	12	AY048733	AY048733 CRIM plas
16	22.4	93.3	4105	12	AF271663	AF271663 Cloning v
17	22.4	93.3	4190	12	XXU13848	U13848 pEXCell Cio
18	22.4	93.3	4549	12	AF178452	AF178452 Integrati
19	22.4	93.3	4549	12	AF178453	AF178453 Integrati
20	22.4	93.3	4782	12	AY048740	AY048740 CRIM plas
21	22.4	93.3	5641	6	AX113748	AX113748 Sequence
22	22.4	93.3	5670	6	AX113749	AX113749 Sequence
23	22.4	93.3	5706	12	AY150267	AY150267 CRIM plas
24	22.4	93.3	5739	12	AY054372	AY054372 CRIM plas
25	22.4	93.3	5771	12	AY150262	AY150262 CRIM plas
26	22.4	93.3	5771	12	AY150263	AY150263 CRIM plas
27	22.4	93.3	5814	12	AY150265	AY150265 CRIM plas
28	22.4	93.3	5826	6	AX113746	AX113746 Sequence
29	22.4	93.3	5911	12	AY150268	AY150268 CRIM plas
30	22.4	93.3	5948	12	AY054373	AY054373 CRIM plas
31	22.4	93.3	5986	12	AY150264	AY150264 CRIM plas
32	22.4	93.3	6000	12	U66308	U66308 Expression
33	22.4	93.3	6043	6	AX370644	AX370644 Sequence
34	22.4	93.3	6071	6	AX113747	AX113747 Sequence
35	22.4	93.3	6200	12	AY150266	AY150266 CRIM plas
36	22.4	93.3	6664	12	AY048713	AY048713 CRIM plas
37	22.4	93.3	6668	12	AY048714	AY048714 CRIM plas
38	22.4	93.3	6742	12	AY048739	AY048739 CRIM plas
39	22.4	93.3	7176	12	AF178451	AF178451 Integrati
40	22.4	93.3	39732	7	AF069529	AF069529 Bacteriop
41	22.4	93.3	42529	12	CVU39284	U39284 Cloning vec
42	22.4	93.3	42530	12	CVU39285	U39285 Cloning vec
43	22.4	93.3	42531	12	CVU39286	U39286 Cloning vec
44	22.4	93.3	42704	12	CVU37692	U37692 Cloning vec
45	22.4	93.3	48502	7	LAMCG	J02459 Bacterioph

ALIGNMENTS

RESULT 1	AX092116	AX092116	Sequence 5 from Patent WO0116345.	243 bp	DNA	linear	PAT 21-MAR-2001
LOCUS	AX092116	Sequence 5 from Patent WO0116345.					
DEFINITION	AX092116	Sequence 5 from Patent WO0116345.					
ACCESSION	AX092116	Sequence 5 from Patent WO0116345.					
VERSION	AX092116.1	GI:13444359					
KEYWORDS							
SOURCE		synthetic construct					
ORGANISM		synthetic construct					
		artificial sequences.					
REFERENCE	1						
AUTHORS		Droege, P.					
TITLE		Sequence-specific dna recombination in eukaryotic cells					
JOURNAL		Patent: WO 0116345-A 5 08-MAR-2001;					
		Droege, Peter (DE)					

FEATURES source Location/Qualifiers  
1..243  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="oligonucleotide"  
BASE COUNT 74 a 34 c 34 g 101 t  
ORIGIN

Query Match 100.0%; Score 24; DB 6; Length 243;  
Best Local Similarity 100.0%; Pred. No. 8.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATCTACTAAGTTG 24  
|||||  
Db 139 GTTCAGCTTTTGATCTACTAAGTTG 162  
|||||

RESULT 2  
LOCUS AR044609 201 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 18 from patent US 5817506.  
ACCESSION AR044609  
VERSION AR044609.1 GI:5966074  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 201)  
AUTHORS Okano,K. and Kanbara,H.  
TITLE Polynucleotide capturing support for capturing, eluting and collecting polynucleotides in a sample solution  
JOURNAL Patent: US 5817506-A 18 OCT-1998;  
FEATURES Location/Qualifiers  
source 1..201  
BASE COUNT 58 a 38 c 27 g 78 t  
ORIGIN

Query Match 93.3%; Score 22.4; DB 6; Length 201;  
Best Local Similarity 95.8%; Pred. No. 42; Mismatches 23; Conservative 0; Indels 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATCTACTAAGTTG 24  
|||||  
Db 40 GTTCAGCTTTTGATCTACTAAGTTG 63  
|||||

RESULT 3  
LOCUS E05439 201 bp DNA linear PAT 29-SEP-1997  
DEFINITION Oligonucleotide.  
ACCESSION E05439  
VERSION E05439.1 GI:2173628  
KEYWORDS JP 1993236997-A/11.  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 201)  
AUTHORS Okano,K. and Kanbara,H.  
TITLE CHIP FOR CATCHING POLYNUCLEOTIDE  
JOURNAL Patent: JP 1993236997-A 11 17-SEP-1993;  
HITACHI LTD  
COMMENT OS Artificial gene  
OC Artificial sequence; Genes.  
PN JP 1993236997-A/11  
PD 17-SEP-1993  
PF 28-FEB-1992 JP 1992042829  
PI OKANO KAZUNOBU, KANBARA HIDEKI  
PC C12Q1/68;  
CC strandedness: Single;  
topology: Linear.  
FEATURES Location/Qualifiers

FEATURES source  
1..201  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 58 a 38 c 27 g 78 t  
ORIGIN

Query Match 93.3%; Score 22.4; DB 6; Length 201;  
Best Local Similarity 95.8%; Pred. No. 42; Mismatches 23; Conservative 0; Indels 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATCTACTAAGTTG 24  
|||||  
Db 40 GTTCAGCTTTTGATCTACTAAGTTG 63  
|||||

RESULT 4  
LOCUS I13139 201 bp DNA linear PAT 26-JUL-1995  
DEFINITION Sequence 18 from patent US 5434049.  
ACCESSION I13139  
VERSION I13139.1 GI:910488  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 201)  
AUTHORS Okano,K. and Kanbara,H.  
TITLE Separation of polynucleotides using supports having a plurality of electrode-containing cells  
JOURNAL Patent: US 5434049-A 18 JUL-1995;  
FEATURES Location/Qualifiers  
source 1..201  
BASE COUNT 58 a 38 c 27 g 78 t  
ORIGIN

Query Match 93.3%; Score 22.4; DB 6; Length 201;  
Best Local Similarity 95.8%; Pred. No. 42; Mismatches 23; Conservative 0; Indels 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATCTACTAAGTTG 24  
|||||  
Db 40 GTTCAGCTTTTGATCTACTAAGTTG 63  
|||||

RESULT 5  
LOCUS I36498 201 bp DNA linear PAT 13-MAY-1997  
DEFINITION Sequence 18 from patent US 5607646.  
ACCESSION I36498  
VERSION I36498.1 GI:2086323  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 201)  
AUTHORS Okano,K. and Kanbara,H.  
TITLE Device for separating polynucleotides having a plurality of electrode-containing cells and movable collecting capillary  
JOURNAL Patent: US 5607646-A 18 04-MAR-1997;  
FEATURES Location/Qualifiers  
source 1..201  
BASE COUNT 58 a 38 c 27 g 78 t  
ORIGIN

Query Match 93.3%; Score 22.4; DB 6; Length 201;  
Best Local Similarity 95.8%; Pred. No. 42; Mismatches 23; Conservative 0; Indels 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATCTACTAAGTTG 24  
|||||  
Db 40 GTTCAGCTTTTGATCTACTAAGTTG 63  
|||||

FEATURES source  
1..201  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 58 a 38 c 27 g 78 t  
ORIGIN

Query Match 93.3%; Score 22.4; DB 6; Length 201;  
Best Local Similarity 95.8%; Pred. No. 42; Mismatches 23; Conservative 0; Indels 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATCTACTAAGTTG 24  
|||||  
Db 40 GTTCAGCTTTTGATCTACTAAGTTG 63  
|||||

Db 40 GTTCAGCTTTTATCTAAGTTG 63

RESULT 6

AX092113

LOCUS AX092113 243 bp DNA linear PAT 21-MAR-2001

DEFINITION Sequence 2 from Patent WO0116345.

ACCESSION AX092113

VERSION AX092113.1 GI:13444356

KEYWORDS Bacteriophage lambda

SOURCE Bacteriophage lambda

ORGANISM Bacteriophage lambda

REFERENCE 1

AUTHORS Droegge, P.

TITLE Sequence-specific dna recombination in eukaryotic cells

JOURNAL Patent: WO 0116345-A 2 08-MAR-2001;

Dröge, Peter (DE)

FEATURES

source

1. .243

/organism="Bacteriophage lambda"

/mol\_type="genomic DNA"

/db\_xref="taxon:10710"

BASE COUNT 74 a 34 c 33 g 102 t

ORIGIN

Query Match 93.3%; Score 22.4; DB 6; Length 243;

Best Local Similarity 95.8%; Pred. No. 41;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTACTAAGTTG 24

Db 139 GTTCAGCTTTTATCTAAGTTG 162

RESULT 7

LAMINTATT/c

LOCUS LAMINTATT 361 bp DNA linear PHG 28-APR-1993

DEFINITION Bacteriophage lambda int gene region.

ACCESSION M23841

VERSION M23841.1 GI:1215177

KEYWORDS site-specific recombination.

SOURCE Bacteriophage lambda

ORGANISM Bacteriophage lambda

REFERENCE 1

AUTHORS Schmeissner U., McKenney K., Rosenberg M. and Court, D.

TITLE Removal of a terminator structure by RNA processing regulates int gene expression

JOURNAL J. Mol. Biol. 176 (1), 39-53 (1984)

MEDLINE 84242838

PUBMED 6234400

COMMENT Original source text: Bacteriophage lambda DNA.

FEATURES

source

1. .361

/organism="Bacteriophage lambda"

/mol\_type="genomic DNA"

/db\_xref="taxon:10710"

BASE COUNT 134 a 60 c 60 g 107 t

ORIGIN

Query Match 93.3%; Score 22.4; DB 7; Length 361;

Best Local Similarity 95.8%; Pred. No. 39;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTACTAAGTTG 24

Db 121 GTTCAGCTTTTATCTAAGTTG 98

RESULT 8

AX101000

LOCUS AX101000 610 bp DNA linear PAT 10-APR-2001

DEFINITION Sequence 1 from Patent WO0121780.

ACCESSION AX101000

VERSION AX101000.1 GI:13619857

KEYWORDS Nicotiana tabacum (common tobacco)

SOURCE Nicotiana tabacum

ORGANISM Nicotiana tabacum

REFERENCE 1

AUTHORS Meyer, P.L. and Zubko, E.L.

TITLE Targeted gene removal

JOURNAL Patent: WO 0121780-A 1 29-MAR-2001;

THE UNIVERSITY OF LEEDS (GB)

FEATURES

source

1. .610

/organism="Nicotiana tabacum"

/mol\_type="genomic DNA"

/db\_xref="taxon:4097"

BASE COUNT 164 a 108 c 110 g 228 t

ORIGIN

Query Match 93.3%; Score 22.4; DB 6; Length 610;

Best Local Similarity 95.8%; Pred. No. 35;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTACTAAGTTG 24

Db 277 GTTCAGCTTTTATCTAAGTTG 300

RESULT 9

MACHSS/c

LOCUS MACHSS 1668 bp mRNA linear PRI 06-FEB-1999

DEFINITION Macaca fascicularis mRNA for hydroxysteroid sulfotransferase subunit, complete cds.

ACCESSION D85521

VERSION D85521.1 GI:1345405

KEYWORDS hydroxysteroid sulfotransferase subunit.

SOURCE Macaca fascicularis (crab-eating macaque)

ORGANISM Macaca fascicularis

REFERENCE 1

AUTHORS Ogura, K., Satsukawa, M., Kato, K., Okuda, H. and Watabe, T.

TITLE Molecular cloning of monkey liver hydroxysteroid sulfotransferase

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1668)

AUTHORS Ogura, K.

TITLE Direct Submission

JOURNAL Submitted (23-MAY-1996) Kenichiro Ogura, Tokyo University of Pharmacy and Life Science, Department of Drug Metabolism and Molecular Toxicology; 1432-1 Horinouchi, Hachioji, Tokyo 192-0392, Japan (E-mail:ogurakeps.toyaku.ac.jp, Tel:+81-426-76-4518, Fax:+81-426-76-4517)

FEATURES

source

1. .1668

/organism="Macaca fascicularis"

/mol\_type="mRNA"

/db\_xref="taxon:9541"

/clone="monHST-1"

/sex="male"

/tissue\_type="liver"

/clone\_lib="lambda gt11"

611. .1468

/codon\_start=1

/product="hydroxysteroid sulfotransferase subunit"

/protein\_id="BAAL2823.1"

/db\_xref="GI:1345406"

/translation="MSDDFLWFEGIAFPNMGFRSETLRKVRDEFVKDEVDVILTYPK"

CDS



```

PVLVFKTDLGSLNELQDEAARLSWLATTGVCFAVLVDVVTAGRDMLLLGEVPGQDL
LSHLLAPAEKVSIADARLRLHLDTPATCPDHOAKHRIERARTMEAGLVDDDLDE
BHOGLAPAEFLARLARVPDGDVLAVTHGDACLNMVYENGFSGFDGRLGVADRY
QDIALATRDIAELGEWADRELVLVGLYGAAPDSQRIARFLDLDRFF"
BASE COUNT      767 a   718 c   716 g   758 t
ORIGIN

Query Match      93.3%; Score 22.4; DB 12; Length 2959;
Best Local Similarity 95.8%; Pred. No. 28;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGTACTAAGTTG 24
|||||
Db 1079 GTTCAGCTTTTGTACTAAGTTG 1102
|||||

RESULT 12
AF178449
LOCUS      AF178449          3485 bp      DNA      linear      SYN 20-JUL-2000
DEFINITION Integration vector pCD11PKS chloramphenicol transacetylase (cat)
and beta-galactosidase alpha peptide (lacZa) genes, complete cds.
ACCESSION AF178449
VERSION    AF178449.1 GI:9294785
KEYWORDS   Integration vector pCD11PKS
SOURCE     Integration vector pCD11PKS
ORGANISM   artificial sequences; vectors.
REFERENCE  1 (bases 1 to 3485)
AUTHORS    Platt,R., Drescher,C., Park,S.K. and Phillips,G.J.
TITLE      Genetic system for reversible integration of DNA constructs and
lacZ gene fusions into the Escherichia coli chromosome
JOURNAL    Plasmid 43 (1), 12-23 (2000)
MEDLINE    20079288
PUBMED     10610816
REFERENCE  2 (bases 1 to 3485)
AUTHORS    Platt,R., Drescher,C., Park,S.K. and Phillips,G.J.
TITLE      Direct Submission
JOURNAL    Submitted (18-AUG-1999) Microbiology, Iowa State University, 207
Science I Building, Ames, IA 50011, USA
FEATURES   Location/Qualifiers
            source
            1..3485
            /organism="Integration vector pCD11PKS"
            /mol_type="genomic DNA"
            /db_xref="taxon:106602"
            /lab_host="Escherichia coli"
            1..385
            rep_origin
            1..385
            /note="R6Kgamma"
            668..915
            misc_feature
            /note="attP; attachment site from bacteriophage lambda"
            complement(1327..1986)
            gene
            /gene="cat"
            complement(1327..1986)
            CDS
            /gene="cat"
            /codon_start=1
            /transl_table=11
            /product="chloramphenicol transacetylase"
            /protein_id="AAF86671.1"
            /db_xref="GI:9294787"
            /translation="MEKLTGYTYVDISQWRKHEPFAQSVQACTYNTQVLDITAF
            LKTVKKNKHKFPFAPHILARLNMNAHPEFRMAKDGELVWDSVHPCYTVFHEQTETP
            SSLWSEYHDDRFQFLHIYSQDVACYNLAYFPKGFIEFMFVSANPMWSTFDLNV
            ANMNDFFAPVFTMGKYTQGDVKVLMPLAIQVHVAVCDGDFHVGRLNELQYCDWEQGG
            A"
            complement(2754..3137)
            gene
            /gene="lacZa"
            complement(2754..3137)
            CDS
            /gene="lacZa"
            /codon_start=1
            /transl_table=11
            /product="beta-galactosidase alpha peptide"
            /protein_id="AAF86670.1"
            /db_xref="GI:9294786"
            /translation="WTMTTPSAQLTLTKGNKSWVPGPSRSTVSIISLNSCSPGDPL

```

```

VLERPPRWSSNPYSESYARSVLAVVLQRDWENPGVTQLNRLAAHPPFASWRNSEE
ASTDRPSQQLRSNGEWLTFPVAH"
2972..3361
/notes="multiple cloning site"
BASE COUNT      968 a   797 c   803 g   917 t
ORIGIN

Query Match      93.3%; Score 22.4; DB 12; Length 3485;
Best Local Similarity 95.8%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGTACTAAGTTG 24
|||||
Db 811 GTTCAGCTTTTGTACTAAGTTG 834
|||||

RESULT 13
AF178450
LOCUS      AF178450          3485 bp      DNA      linear      SYN 20-JUL-2000
DEFINITION Integration vector pCD11PSK chloramphenicol transacetylase (cat)
and beta-galactosidase alpha peptide (lacZa) genes, complete cds.
ACCESSION AF178450
VERSION    AF178450.1 GI:9294788
KEYWORDS   Integration vector pCD11PSK
SOURCE     Integration vector pCD11PSK
ORGANISM   artificial sequences; vectors.
REFERENCE  1 (bases 1 to 3485)
AUTHORS    Platt,R., Drescher,C., Park,S.K. and Phillips,G.J.
TITLE      Genetic system for reversible integration of DNA constructs and
lacZ gene fusions into the Escherichia coli chromosome
JOURNAL    Plasmid 43 (1), 12-23 (2000)
MEDLINE    20079288
PUBMED     10610816
REFERENCE  2 (bases 1 to 3485)
AUTHORS    Platt,R., Drescher,C., Park,S.K. and Phillips,G.J.
TITLE      Direct Submission
JOURNAL    Submitted (18-AUG-1999) Microbiology, Iowa State University, 207
Science I Building, Ames, IA 50011, USA
FEATURES   Location/Qualifiers
            source
            1..3485
            /organism="Integration vector pCD11PSK"
            /mol_type="genomic DNA"
            /db_xref="taxon:106603"
            /lab_host="Escherichia coli"
            1..385
            rep_origin
            1..385
            /note="R6Kgamma"
            668..915
            misc_feature
            /note="attP; attachment site from bacteriophage lambda"
            complement(1327..1986)
            gene
            /gene="cat"
            complement(1327..1986)
            CDS
            /gene="cat"
            /codon_start=1
            /transl_table=11
            /product="chloramphenicol transacetylase"
            /protein_id="AAF86673.1"
            /db_xref="GI:9294790"
            /translation="MEKLTGYTYVDISQWRKHEPFAQSVQACTYNTQVLDITAF
            LKTVKKNKHKFPFAPHILARLNMNAHPEFRMAKDGELVWDSVHPCYTVFHEQTETP
            SSLWSEYHDDRFQFLHIYSQDVACYNLAYFPKGFIEFMFVSANPMWSTFDLNV
            ANMNDFFAPVFTMGKYTQGDVKVLMPLAIQVHVAVCDGDFHVGRLNELQYCDWEQGG
            A"
            complement(2754..3137)
            gene
            /gene="lacZa"
            complement(2754..3137)
            CDS
            /gene="lacZa"
            /codon_start=1
            /transl_table=11
            /product="beta-galactosidase alpha peptide"
            /protein_id="AAF86672.1"
            /db_xref="GI:9294789"
            /translation="WTMTTPSAQLTLTKGNKSWSTAVAAALELVDPPGCRNSISLS

```



```
IPSTRSGGPVNSPYSBSYSLAVLVQRDWNPGVTQNLRLAAHPFASWRNGEE
ARTDRPSQQLRSLNGEWLTPVAAH"
misc_feature 2972..3081
/note="multiple cloning site"
BASE COUNT 967 a 798 c 802 g 918 t
ORIGIN
Query Match 93.3%; Score 22.4; DB 12; Length 3485;
Best Local Similarity 95.8%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GTTCAGCTTTTGTACTAAGTTG 24
|||||
Db 811 GTTCAGCTTTTGTACTAAGTTG 834
|||||

RESULT 14
AY048716 3663 bp DNA circular SYN 17-OCT-2001
LOCUS
DEFINITION CRIM plasmid pAH63, complete sequence.
ACCESSION AY048716
VERSION AY048716.1 GI:16209075
KEYWORDS
SOURCE CRIM plasmid pAH63
ORGANISM CRIM plasmid pAH63
REFERENCE 1 (bases 1 to 3663)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Conditional-replication, integration, excision, and retrieval
plasmid-host systems for gene structure-function studies of
bacteria
JOURNAL J. Bacteriol. 183 (21), 6384-6393 (2001)
MEDLINE 21475783
PUBMED 11591683
REFERENCE 2 (bases 1 to 3663)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2001) Biological Sciences, Purdue University,
Lilly Hall, West Lafayette, IN 47907, USA
FEATURES
source
1..3663
/organism="CRIM plasmid pAH63"
/mol_type="genomic DNA"
/db_xref="taxon:172341"
101..1030
/gene="pstS"
101..1030
/gene="pstS"
/codon_start=1
/transl_table=11
/product="pstS"
/protein_id="AAL09852.1"
/db_xref="GI:16209076"
/translation="MKVNRVTAVVAATLSMAFVSFAEASLTGAGATFPAPVYAKW
ADTYQKGTGNKVNGTQGGSGGVQIIANTVDFGASDAPLSDEKLAQEGLFQPTVIG
GLVAVNAPVLKSGELVDGKTDLYLGIKKWDDAIAKLNPLKFSQNTAVVR
ADGSGTSFVTSYLAKEWENKNNVGTGTPKPIGLGKNGDIAAFVORLPQATGY
VEYAYAKNNLAYTKLISADGKPVSPTEENFANAAGADWSKTFQADLTNKGEDAWP
ITSTFTLLHKQKPEQDRSAEILRLGVQNRG"
complement(2423..3217)
/codon_start=1
/transl_table=11
/product="kan"
/protein_id="AAL09853.1"
/db_xref="GI:16209077"
/translation="MIEQDGLHAGSPAANVERLFGYDWAQQTICGSDAAVFLSAGQR
PVLVFKTDLGALNEQDEAARLSWLTATGVPCAAVLDVVTEAGRDWLLIGVEFGQDL
LSSHAPAEKVSINADAMRLHTLDPATCFPHQAKHRIERARTMEAGLVDDDLDE
EHQGLAPAEFLPARLKAMPDGEDLVVTHGDACLNINWENGFRSGFIDCGRLGVADRY
QDIALATEDIAEELGGWADRFLVLYGIAAPDSQRIAFYRLLDFF"
complement(3174..3863)
/promoter
/note="Psyml; from pHH7013"
BASE COUNT 954 a 892 c 914 g 903 t

IPSTRSGGPVNSPYSBSYSLAVLVQRDWNPGVTQNLRLAAHPFASWRNGEE
ARTDRPSQQLRSLNGEWLTPVAAH"
misc_feature 2972..3081
/note="multiple cloning site"
BASE COUNT 967 a 798 c 802 g 918 t
ORIGIN
Query Match 93.3%; Score 22.4; DB 12; Length 3485;
Best Local Similarity 95.8%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GTTCAGCTTTTGTACTAAGTTG 24
|||||
Db 811 GTTCAGCTTTTGTACTAAGTTG 834
|||||

RESULT 14
AY048716 3663 bp DNA circular SYN 17-OCT-2001
LOCUS
DEFINITION CRIM plasmid pAH63, complete sequence.
ACCESSION AY048716
VERSION AY048716.1 GI:16209075
KEYWORDS
SOURCE CRIM plasmid pAH63
ORGANISM CRIM plasmid pAH63
REFERENCE 1 (bases 1 to 3663)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Conditional-replication, integration, excision, and retrieval
plasmid-host systems for gene structure-function studies of
bacteria
JOURNAL J. Bacteriol. 183 (21), 6384-6393 (2001)
MEDLINE 21475783
PUBMED 11591683
REFERENCE 2 (bases 1 to 3663)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2001) Biological Sciences, Purdue University,
Lilly Hall, West Lafayette, IN 47907, USA
FEATURES
source
1..3663
/organism="CRIM plasmid pAH63"
/mol_type="genomic DNA"
/db_xref="taxon:172341"
101..1030
/gene="pstS"
101..1030
/gene="pstS"
/codon_start=1
/transl_table=11
/product="pstS"
/protein_id="AAL09852.1"
/db_xref="GI:16209076"
/translation="MKVNRVTAVVAATLSMAFVSFAEASLTGAGATFPAPVYAKW
ADTYQKGTGNKVNGTQGGSGGVQIIANTVDFGASDAPLSDEKLAQEGLFQPTVIG
GLVAVNAPVLKSGELVDGKTDLYLGIKKWDDAIAKLNPLKFSQNTAVVR
ADGSGTSFVTSYLAKEWENKNNVGTGTPKPIGLGKNGDIAAFVORLPQATGY
VEYAYAKNNLAYTKLISADGKPVSPTEENFANAAGADWSKTFQADLTNKGEDAWP
ITSTFTLLHKQKPEQDRSAEILRLGVQNRG"
complement(2423..3217)
/codon_start=1
/transl_table=11
/product="kan"
/protein_id="AAL09853.1"
/db_xref="GI:16209077"
/translation="MIEQDGLHAGSPAANVERLFGYDWAQQTICGSDAAVFLSAGQR
PVLVFKTDLGALNEQDEAARLSWLTATGVPCAAVLDVVTEAGRDWLLIGVEFGQDL
LSSHAPAEKVSINADAMRLHTLDPATCFPHQAKHRIERARTMEAGLVDDDLDE
EHQGLAPAEFLPARLKAMPDGEDLVVTHGDACLNINWENGFRSGFIDCGRLGVADRY
QDIALATEDIAEELGGWADRFLVLYGIAAPDSQRIAFYRLLDFF"
complement(3174..3863)
/promoter
/note="Psyml; from pHH7013"
BASE COUNT 954 a 892 c 914 g 903 t

IPSTRSGGPVNSPYSBSYSLAVLVQRDWNPGVTQNLRLAAHPFASWRNGEE
ARTDRPSQQLRSLNGEWLTPVAAH"
misc_feature 2972..3081
/note="multiple cloning site"
BASE COUNT 967 a 798 c 802 g 918 t
ORIGIN
Query Match 93.3%; Score 22.4; DB 12; Length 3485;
Best Local Similarity 95.8%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GTTCAGCTTTTGTACTAAGTTG 24
|||||
Db 1783 GTTCAGCTTTTGTACTAAGTTG 1806
|||||

RESULT 15
AY048733 3695 bp DNA circular SYN 17-OCT-2001
LOCUS
DEFINITION CRIM plasmid pAH150, complete sequence.
ACCESSION AY048733
VERSION AY048733.1 GI:16209150
KEYWORDS
SOURCE CRIM plasmid pAH150
ORGANISM CRIM plasmid pAH150
REFERENCE 1 (bases 1 to 3695)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Conditional-replication, integration, excision, and retrieval
plasmid-host systems for gene structure-function studies of
bacteria
JOURNAL J. Bacteriol. 183 (21), 6384-6393 (2001)
MEDLINE 21475783
PUBMED 11591683
REFERENCE 2 (bases 1 to 3695)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-2001) Biological Sciences, Purdue University,
Lilly Hall, West Lafayette, IN 47907, USA
FEATURES
source
1..3695
/organism="CRIM plasmid pAH150"
/mol_type="genomic DNA"
/db_xref="taxon:172358"
complement(38..193)
/note="araC' terminal fragment"
198..507
/note="araB3p"
548..1237
/codon_start=1
/transl_table=11
/product="phoB"
/protein_id="AAL09901.1"
/db_xref="GI:16209152"
/translation="MARRILVDEAPIREMVCFLVQNGFPVEAEYDSAVNQLNE
PWPDLILLDWLPGSGGQIFKHLKRSMTDRDIPWMLTARGEEDRVRGLETGADDY
ITKPSPKELVARIKAVMRISPMAVEVIEWQGLSDPTSHRWAGEEPLMGPTER
KLHFFFWTHPERSVYREGQLNHVGTNVYVEDRTVDVHRLRLKALEPFGGHDVWQTV
RGTGYRFSTRF"
complement(2455..3249)
/codon_start=1
/transl_table=11
/product="kan"
/protein_id="AAL09900.1"
/db_xref="GI:16209151"
/translation="MIEQDGLHAGSPAANVERLFGYDWAQQTICGSDAAVFLSAGQR
PVLVFKTDLGALNEQDEAARLSWLTATGVPCAAVLDVVTEAGRDWLLIGVEFGQDL
LSSHAPAEKVSINADAMRLHTLDPATCFPHQAKHRIERARTMEAGLVDDDLDE
EHQGLAPAEFLPARLKAMPDGEDLVVTHGDACLNINWENGFRSGFIDCGRLGVADRY
QDIALATEDIAEELGGWADRFLVLYGIAAPDSQRIAFYRLLDFF"
complement(914 c 913 g 919 t
BASE COUNT 949 a 914 c 913 g 919 t
ORIGIN
Query Match 93.3%; Score 22.4; DB 12; Length 3695;
Best Local Similarity 95.8%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GTTCAGCTTTTGTACTAAGTTG 24
```

Db 1815 GTTCAGCTTTTATATAAGTTG 1838

Search completed: December 3, 2003, 13:18:42  
Job time : 653.562 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 1076.23 Seconds  
(without alignments)  
541.991 Million cell updates/sec

Title: US-10-082-772-3

Perfect score: 24

Sequence: 1 gttcagctttttgataactaattg 24

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estm:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estic:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	23	95.8	664	28	AQ991011
C 2	22.4	93.3	206	13	BQ156416
C 3	22.4	93.3	299	13	BY115594
C 4	22.4	93.3	472	13	BQ157398

C 5	22.4	93.3	473	13	BQ156404
C 6	22.4	93.3	597	12	BI422679
C 7	22.4	93.3	695	28	AQ991039
C 8	22.4	93.3	712	28	AQ990809
C 9	22.4	93.3	743	28	AQ990346
C 10	22.4	93.3	764	28	AQ990110
C 11	22.4	93.3	769	28	AQ990470
C 12	21.4	89.2	395	28	AQ991303
C 13	21.4	89.2	751	28	AQ989566
C 14	20.8	86.7	672	28	AQ990864
C 15	20.8	86.7	753	28	AQ990861
C 16	20.8	86.7	770	28	AQ991774
C 17	20.8	86.7	791	28	AQ991791
C 18	20.8	86.7	808	28	AQ990388
C 19	19.8	82.5	719	28	AQ991352
C 20	19.8	82.5	756	28	AQ991732
C 21	19.8	82.5	787	29	BZ971242
C 22	19.2	80.0	220	9	AI856003
C 23	19.2	80.0	675	28	AQ991241
C 24	19.2	80.0	708	28	AQ990869
C 25	19.2	80.0	746	28	AQ990173
C 26	18.8	78.3	530	28	AZ263252
C 27	18.8	78.3	559	28	AZ300017
C 28	18.8	78.3	653	14	CB422481
C 29	18.8	78.3	676	29	AG147616
C 30	18.8	78.3	731	28	AQ975703
C 31	18.8	78.3	891	29	AG093904
C 32	18.2	75.8	259	10	BG156704
C 33	18.2	75.8	388	13	BQ697293
C 34	18.2	75.8	389	9	AW758911
C 35	18.2	75.8	391	13	BY510914
C 36	18.2	75.8	403	9	AW153070
C 37	18.2	75.8	404	13	BU493456
C 38	18.2	75.8	407	10	BG317764
C 39	18.2	75.8	428	9	AI449436
C 40	18.2	75.8	441	9	AI510342
C 41	18.2	75.8	449	28	AZ413832
C 42	18.2	75.8	452	9	AA120033
C 43	18.2	75.8	472	12	BM188477
C 44	18.2	75.8	489	29	AG256781
C 45	18.2	75.8	500	28	BH885061

ALIGNMENTS

RESULT 1	AQ991011/C	AQ991011	684 bp	DNA	linear	GSS 14-AUG-2000
LOCUS	AQ991011/C	Ric01864	Photobacterium luminescens strain W14 M13 library			
DEFINITION	AQ991011	Photobacterium luminescens genomic clone PLG01864, genomic survey sequence.				
ACCESSION	AQ991011	AQ991011.1	GI:9649605			
VERSION	AQ991011	GSS.				
KEYWORDS	AQ991011	Photobacterium luminescens				
SOURCE	AQ991011	Photobacterium luminescens				
ORGANISM	AQ991011	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photobacterium				
REFERENCE	AQ991011	1 (bases 1 to 684)				
AUTHORS	AQ991011	ifrench-Constant R.H., Waterfield N., Burland V., Perna N.T., Daborn P.J., Bower D. and Blattner F.R.				
TITLE	AQ991011	A genomic sample sequence of the entomopathogenic bacterium Photobacterium luminescens W14: potential implications for virulence				
JOURNAL	AQ991011	Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)				
MEDLINE	AQ991011	20378633				
PUBMED	AQ991011	10919786				
COMMENT	AQ991011	Contact: ifrench-Constant RH Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779				

Email: bssrfcc@bath.ac.uk  
 This is one of 2,122 random reads from the M13 library. For  
 annotation of identified clones (BLASTX, BLASTN and mapping to E.  
 coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic  
 Acids Res.  
 Seq primer: M13 Forward  
 Class: shotgun.

## FEATURES

source  
 1. .664  
 /organism="Photorhabdus luminescens"  
 /mol\_type="genomic DNA"  
 /strain="W14"  
 /db\_xref="taxon:29488"  
 /clone="PLG01864"  
 /dev\_stage="primary phase variant"  
 /clone\_lib="Photorhabdus luminescens strain W14 M13  
 library"  
 /notes="Genomic DNA from strain W14 was size selected (1-2  
 kb) and then cloned into M13 Janus."

BASE COUNT 178 a 146 c 162 g 175 t 3 others  
 ORIGIN

Query Match 95.8%; Score 23; DB 28; Length 664;  
 Best Local Similarity 95.8%; Pred. No. 30;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTACTAAGTTG 24  
 |||||  
 Db 638 GTTCAGCTTTTGTACTAAGTTG 615

## RESULT 2

BQ156416/c  
 LOCUS BQ156416 206 bp mRNA linear EST 24-APR-2002  
 DEFINITION NF092F02IR1F1027 Irradiated Medicago truncatula cDNA clone  
 NF092F02IR 5', mRNA sequence.

ACCESSION BQ156416  
 VERSION BQ156416.1 GI:20293475  
 KEYWORDS EST.  
 SOURCE Medicago truncatula (barrel medic)  
 ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

1 (bases 1 to 206)

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,  
 Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula irradiated library  
 Unpublished

Contact: May GD

Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2310 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7391

Fax: 580 221 7380

Email: gdmay@noble.org

Insert Length: 206 Std Error: 0.00

Plate: 032 row: F column: 02

Seq primer: TCACACGGAACAGCTATGAC.

## FEATURES

source  
 1. .206  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3880"  
 /clone="NF092F02IR"  
 /tissue\_type="seedlings"  
 /dev\_stage="seedling"  
 /clone\_lib="irradiated"

/notes="Vector: Lambda zap; Seedlings were exposed either  
 to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.  
 Gamma-irradiated samples were harvested at 6, 12, 24 and

48 hours after treatment. UV-irradiated samples were  
 harvested 24 hours post-treatment. cDNA was prepared from  
 polyA+ enriched, pooled samples of equivalent amounts of  
 total RNA from each sample. The cDNA was directionally  
 ligated into the Uni-zap XR vector (Stratagene) and  
 packaged using the Gigapack III Gold packaging extracts.  
 Phagemids containing cDNA inserts were in vivo excised  
 from the recombinant Uni-zap XR vector using ExAssist  
 helper phage and the E. coli strain XL1-Blue MRP,  
 (Stratagene). Excised plasmids were plated using SOLR  
 cells."

BASE COUNT 81 a 27 c 39 g 59 t

Query Match 93.3%; Score 22.4; DB 13; Length 206;  
 Best Local Similarity 95.8%; Pred. No. 37;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTACTAAGTTG 24  
 |||||  
 Db 167 GTTCAGCTTTTGTACTAAGTTG 144

## RESULT 3

BY115594  
 LOCUS BY115594 299 bp mRNA linear EST 08-DEC-2002  
 DEFINITION BY115594 RIKEN full-length enriched, 18 days embryo whole body Mus  
 musculus cDNA clone L430040C03 5', mRNA sequence.

ACCESSION BY115594

VERSION BY115594.1 GI:26226695

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 299)

## REFERENCE

## AUTHORS

Okazaki, Y., Furuno, M., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,  
 Nikaide, I., Osato, N., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,  
 Gojobori, T., Baldarelli, R., Hill, D.P., Buit, C., Hume, D.A.,  
 Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,  
 Beisel, K.W., Blake, J.A., Bradt, D., Brusci, V., Chothia, C., Corbani,  
 L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,  
 A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,  
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,  
 Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,  
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,  
 P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,  
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,  
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,  
 Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,  
 B.Z., Ringwald, M., Sander, A., Schneider, C., Semple, C.A., Setou,  
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,  
 R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,  
 Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yamanaka,  
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, X., Zimmer, A.,  
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,  
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,  
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,  
 K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,  
 E.S., Rogers, J., Birney, E. and Hayashizaki, Y.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute Of Physical and Chemical Research (RIKEN)  
 1-7-22 Suenho-cho, Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222

Fax: 81-45-503-9216  
 Email: genome-res@gsr.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane  
 T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,  
 Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakaki, K.,  
 Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,  
 M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct  
 Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in Riken.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

#### FEATURES

Location/Qualifiers

1. 299  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="L430040C03"  
 /tissue\_type="whole body"  
 /dev\_stage="18 days embryo"  
 /clone\_lib="RIKEN full-length enriched, 18 days embryo  
 whole body"

85 a 50 c 54 g 110 t

#### BASE COUNT

ORIGIN

Query Match 93.3%; Score 22.4; DB 13; Length 299;  
 Best Local Similarity 95.8%; Pred. No. 42;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATCTACTAGTTG 24

Db 246 GTTCAGCTTTTGATCTACTAGTTG 269

#### RESULT 4

LOCUS

DEFINITION BQ157398 472 bp mRNA linear EST 24-APR-2002  
 NF1040071R1F1062 Irradiated Medicago truncatula cDNA clone

ACCESSION BQ157398

VERSION BQ157398.1 GI:20294457

KEYWORDS

SOURCE Medicago truncatula (barrel medic)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

REFERENCE

AUTHORS

1 (bases 1 to 472)  
 Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,  
 Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

TITLE

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL

COMMENT

Medicago truncatula irradiated library

Unpublished

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7391  
 Fax: 580 221 7380  
 Email: gdmay@noble.org

Insert Length: 472 Std Error: 0.00

Plate: 104 row: D column: 07

Seq primer: TCACACAGGAAACAGCTATGAC.

#### FEATURES

source

Location/Qualifiers

1. 472

/organism="Medicago truncatula"

/mol\_type="mRNA"

/db\_xref="taxon:3880"

/clone="NF1040071R"

/tissue\_type="seedlings"

/dev\_stage="seedling"

/clone\_lib="Irradiated"

/note="Vector: Lambda Zap; Seedlings were exposed either

to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.

Gamma-irradiated samples were harvested at 6, 12, 24 and

48 hours after treatment. UV-irradiated samples were

harvested 24 hours post-treatment. cDNA was prepared from

polyA+ enriched, pooled samples of equivalent amounts of

total RNA from each sample. The cDNA was directionally

ligated into the Uni-Zap XR vector (Stratagene) and

packaged using the Gigapack III Gold packaging extracts.

Phagmids containing cDNA inserts were in vivo excised

from the recombinant Uni-Zap XR vector using EXassist

helper phage and the E. coli strain XL1-Blue MRF'

(Stratagene). Excised plasmids were plated using SOLR

cells."

BASE COUNT 162 a 89 c 95 g 126 t

ORIGIN

Query Match 93.3%; Score 22.4; DB 13; Length 472;

Best Local Similarity 95.8%; Pred. No. 48;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATCTACTAGTTG 24

Db 380 GTTCAGCTTTTGATCTACTAGTTG 357

#### RESULT 5

LOCUS

DEFINITION BQ156404/c

ACCESSION BQ156404

VERSION BQ156404.1 GI:20293463

KEYWORDS

SOURCE Medicago truncatula (barrel medic)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

REFERENCE

AUTHORS

1 (bases 1 to 473)

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,

Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

TITLE

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL

COMMENT

Medicago truncatula irradiated library

Unpublished

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

```

source
1. .473
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NPO92E031R"
/tissue_type="seedlings"
/dev_stage="seedling"
/clone_lib="Irradiated"
/notes="Vector: Lambda Zap; Seedlings were exposed either
to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
Gamma-irradiated samples were harvested at 6, 12, 24 and
48 hours after treatment. UV-irradiated samples were
harvested 24 hours post-treatment. cDNA was prepared from
polyA+ enriched, pooled samples of equivalent amounts of
total RNA from each sample. The cDNA was directionally
ligated into the Uni-Zap XR vector (Stratagene) and
packaged using the Gigapack III Gold packaging extracts.
Phagemids containing cDNA inserts were in vivo excised
from the recombinant Uni-Zap XR vector using Exbaist
helper phage and the E. coli strain XL1-Blue MRF'
(Stratagene). Excised plasmids were plated using SOLR
cells."
BASE COUNT      162 a  90 c  95 g  126 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 13; Length 473;
Best Local Similarity 95.8%; Pred. No. 48;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGCATCAAGTTG 24
|||||
Db 381 GTTCAGCTTTTATCAAGTTG 358

RESULT 6
BI422679/c
LOCUS      BI422679          597 bp      mRNA      linear      EST 16-AUG-2001
DEFINITION EST533345 tomato callus, TAMU Lycopersicon esculentum cDNA clone
cLEC71G2 5' end, mRNA sequence.
ACCESSION      BI422679
VERSION        BI422679.1  GI:15197297
KEYWORDS       EST.
SOURCE          Lycopersicon esculentum (tomato)
ORGANISM       Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE      1 (bases 1 to 597)
AUTHORS        Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Renning
,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato callus tissue
Unpublished
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1. .597
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA96"
/db_xref="taxon:4081"
/clone="cLEC71G2"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XL1-Blue MRF'"
/clone_lib="tomato callus, TAMU"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.

```

```

BASE COUNT      193 a  109 c  131 g  164 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 12; Length 597;
Best Local Similarity 95.8%; Pred. No. 51;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGCATCAAGTTG 24
|||||
Db 552 GTTCAGCTTTTATCAAGTTG 529

RESULT 7
AQ991039/c
LOCUS      AQ991039          695 bp      DNA      linear      GSS 14-AUG-2000
DEFINITION RfC01894 Photorhabdus luminescens strain W14 M13 library
Photorhabdus luminescens genomic clone PLG01894, genomic survey
sequence.
ACCESSION      AQ991039
VERSION        AQ991039.1  GI:9649633
KEYWORDS       GSS.
SOURCE          Photorhabdus luminescens
ORGANISM       Photorhabdus luminescens
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Photorhabdus.
REFERENCE      1 (bases 1 to 695)
AUTHORS        ffrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,
Daborn,P.J., Bowen,D. and Blattner,F.R.
A genomic sample sequence of the entomopathogenic bacterium
Photorhabdus luminescens W14: potential implications for virulence
Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
20378633
PUBMED
10919786
Contact: ffrench-Constant RH
Department of Biology and Biochemistry
University of Bath
South Building, Bath BA2 7AY, UK
Tel: (44) 1225 826621
Fax: (44) 1225 826779
Email: bsr1c@bath.ac.uk
This is one of 2,122 random reads from the M13 library. For
annotation of identified clones (BLASTX, BLASTN and mapping to E.
coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic
Acids Res.
Seq primer: M13 Forward
Class: shotgun.
Location/Qualifiers
1. .695
/organism="Photorhabdus luminescens"
/mol_type="genomic DNA"
/strain="W14"
/db_xref="taxon:29488"
/clone="PLG01894"
/dev_stage="primary phase variant"
/clone_lib="Photorhabdus luminescens strain W14 M13
library"
/notes="Genomic DNA from strain W14 was size selected (1-2
kb) and then cloned into M13 Janus."
BASE COUNT      193 a  148 c  165 g  187 t  2 others
ORIGIN
Query Match      93.3%; Score 22.4; DB 28; Length 695;
Best Local Similarity 95.8%; Pred. No. 54;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGCATCAAGTTG 24
|||||
Db 639 GTTCAGCTTTTATCAAGTTG 616

RESULT 8

```

Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST Library"

AQ990809/c  
LOCUS AQ990809 712 bp DNA linear GSS 14-AUG-2000  
DEFINITION RfC01638 Photorhabdus luminescens strain W14 M13 library  
Photorhabdus luminescens genomic clone PLG01638, genomic survey  
sequence.  
ACCESSION AQ990809 GI:9649403  
VERSION AQ990809.1  
KEYWORDS GSS.  
SOURCE Photorhabdus luminescens  
ORGANISM Photorhabdus luminescens  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Photorhabdus.  
REFERENCE 1 (bases 1 to 712)  
AUTHORS ffrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,  
Daborn,P.J., Bowen,D. and Blattner,F.R.  
TITLE A genomic sample sequence of the entomopathogenic bacterium  
Photorhabdus luminescens W14: potential implications for virulence  
Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)  
JOURNAL  
MEDLINE  
PUBMED 20378633  
COMMENT Contact: ffrench-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK  
Tel: (44) 1225 826621  
Fax: (44) 1225 826779  
Email: bssrfc@bath.ac.uk  
This is one of 2,122 random reads from the M13 library. For  
annotation of identified clones (BLASTX, BLASTN and mapping to E.  
coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic  
Acids Res.  
Seq primer: M13 Forward  
Class: shotgun.  
FEATURES  
source  
1..712  
/organism="Photorhabdus luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG01638"  
/dev\_stage="primary phase variant"  
/clone\_lib="Photorhabdus luminescens strain W14 M13  
library"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."  
BASE COUNT 218 a 144 c 163 g 187 t  
ORIGIN  
Query Match 93.3%; Score 22.4; DB 28; Length 712;  
Best Local Similarity 95.8%; Pred. No. 54;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||  
Db 453 GTTCAGCTTTTGTACTAAGTTG 430  
|||||  
RESULT 9  
AQ990346/c  
LOCUS AQ990346 743 bp DNA linear GSS 14-AUG-2000  
DEFINITION RfC01106 Photorhabdus luminescens strain W14 M13 library  
Photorhabdus luminescens genomic clone PLG01106, genomic survey  
sequence.  
ACCESSION AQ990346 GI:9648940  
VERSION AQ990346  
KEYWORDS GSS.  
SOURCE Photorhabdus luminescens  
ORGANISM Photorhabdus luminescens  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Photorhabdus.  
REFERENCE 1 (bases 1 to 743)  
AUTHORS ffrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,  
Daborn,P.J., Bowen,D. and Blattner,F.R.

TITLE A genomic sample sequence of the entomopathogenic bacterium  
Photorhabdus luminescens W14: potential implications for virulence  
Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)  
JOURNAL  
MEDLINE  
PUBMED 20378633  
COMMENT Contact: ffrench-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK  
Tel: (44) 1225 826621  
Fax: (44) 1225 826779  
Email: bssrfc@bath.ac.uk  
This is one of 2,122 random reads from the M13 library. For  
annotation of identified clones (BLASTX, BLASTN and mapping to E.  
coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic  
Acids Res.  
Seq primer: M13 Forward  
Class: shotgun.  
FEATURES  
source  
1..743  
/organism="Photorhabdus luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG01106"  
/dev\_stage="primary phase variant"  
/clone\_lib="Photorhabdus luminescens strain W14 M13  
library"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."  
BASE COUNT 214 a 158 c 169 g 200 t  
ORIGIN  
Query Match 93.3%; Score 22.4; DB 28; Length 743;  
Best Local Similarity 95.8%; Pred. No. 55;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||  
Db 626 GTTCAGCTTTTGTACTAAGTTG 603  
|||||  
RESULT 10  
AQ990110/c  
LOCUS AQ990110 764 bp DNA linear GSS 14-AUG-2000  
DEFINITION RfC00827 Photorhabdus luminescens strain W14 M13 library  
Photorhabdus luminescens genomic clone PLG00827, genomic survey  
sequence.  
ACCESSION AQ990110 GI:9648704  
VERSION AQ990110.1  
KEYWORDS GSS.  
SOURCE Photorhabdus luminescens  
ORGANISM Photorhabdus luminescens  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Photorhabdus.  
REFERENCE 1 (bases 1 to 764)  
AUTHORS ffrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,  
Daborn,P.J., Bowen,D. and Blattner,F.R.  
TITLE A genomic sample sequence of the entomopathogenic bacterium  
Photorhabdus luminescens W14: potential implications for virulence  
Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)  
JOURNAL  
MEDLINE  
PUBMED 20378633  
COMMENT Contact: ffrench-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK  
Tel: (44) 1225 826621  
Fax: (44) 1225 826779  
Email: bssrfc@bath.ac.uk  
This is one of 2,122 random reads from the M13 library. For  
annotation of identified clones (BLASTX, BLASTN and mapping to E.  
coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic

```

Acids Res.
Seq primer: M13 Forward
Class: shotgun.
FEATURES
    source            1..764
                        /organism="Photorhabdus luminescens"
                        /mol_type="genomic DNA"
                        /strain="W14"
                        /db_xref="taxon:29488"
                        /clone="PLG00827"
                        /dev_stage="primary phase variant"
                        /clone_lib="Photorhabdus luminescens strain W14 M13
                        library"
BASE COUNT      215 a   170 c   171 g   203 t   5 others
ORIGIN
    Query Match      93.3%; Score 22.4; DB 28; Length 764;
    Best local Similarity 95.8%; Pred. No. 55;
    Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 GTTCAGCTTTTGATCTAGCTTG 24
Db      721 GTTCAGCTTTTATCTAGCTTG 698

RESULT 11
AQ990470/c
LOCUS
DEFINITION
    Photorhabdus luminescens genomic clone PLG01245, genomic survey
    sequence.
ACCESSION
    AQ990470
VERSION
    AQ990470.1 GI:9649064
KEYWORDS
    GSS.
SOURCE
    ORGANISM
        Photorhabdus luminescens
        Photorhabdus luminescens
        Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
        Enterobacteriaceae; Photorhabdus.
        1 (bases 1 to 769)
        Ifrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,
        Daborn,P.J., Bowen,D. and Blattner,F.R.
        A genomic sample sequence of the entomopathogenic bacterium
        Photorhabdus luminescens W14: potential implications for virulence
        Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
        20378633
        10919786
    Contact: french-Constant RH
    Department of Biology and Biochemistry
    University of Bath
    South Building, Bath BA2 7AY, UK
    Tel: (44) 1225 826621
    Fax: (44) 1225 826779
    Email: bsrffc@bath.ac.uk
    Title
        This is one of 2,122 random reads from the M13 library. For
        annotation of identified clones (BLASTX, BLASTN and mapping to E.
        coli K12 Genome) please see french-Constant et al. 2000, Nucleic
        Acids Res.
    Seq primer: M13 Forward
    Class: shotgun.
FEATURES
    source            1..769
                        /organism="Photorhabdus luminescens"
                        /mol_type="genomic DNA"
                        /strain="W14"
                        /db_xref="taxon:29488"
                        /clone="PLG01245"
                        /dev_stage="primary phase variant"
                        /clone_lib="Photorhabdus luminescens strain W14 M13
                        library"
        /note="Genomic DNA from strain W14 was size selected (1-2
        kb) and then cloned into M13 Janus."
        /note="Genomic DNA from strain W14 was size selected (1-2
        kb) and then cloned into M13 Janus."

```



DEFINITION RfC00126 Photorhabdus luminescens strain W14 M13 library  
 Photorhabdus luminescens genomic clone PLG00126, genomic survey  
 sequence.  
 ACCESSION AQ989566  
 VERSION AQ989566.1 GI:9648160  
 KEYWORDS GSS.  
 SOURCE Photorhabdus luminescens  
 ORGANISM Photorhabdus luminescens  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Photorhabdus.  
 REFERENCE 1 (bases 1 to 751)  
 AUTHORS ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,  
 Daborn, P.J., Bowen, D. and Blattner, F.R.  
 TITLE A genomic sample sequence of the entomopathogenic bacterium  
 Photorhabdus luminescens W14: potential implications for virulence  
 JOURNAL Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)  
 MEDLINE 20378633  
 PUBMED 10919786  
 COMMENT Contact: ffrench-Constant RH  
 Department of Biology and Biochemistry  
 University of Bath  
 South Building, Bath BA2 7AY, UK  
 Tel: (44) 1225 826621  
 Fax: (44) 1225 826779  
 Email: bsarfcbath.ac.uk  
 This is one of 2,122 random reads from the M13 library. For  
 annotation of identified clones (BLASTX, BLASTN and mapping to E.  
 coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic  
 Acids Res.  
 Seq primer: M13 Forward  
 Class: shotgun.

FEATURES  
 source  
 Location/Qualifiers  
 1..751  
 /organism="Photorhabdus luminescens"  
 /mol\_type="genomic DNA"  
 /strain="W14"  
 /db\_xref="taxon:29488"  
 /clone="PLG00126"  
 /dev\_stage="primary phase variant"  
 /clone\_lib="Photorhabdus luminescens strain W14 M13  
 library"  
 /note="Genomic DNA from strain W14 was size selected (1-2  
 kb) and then cloned into M13 Janus."  
 BASE COUNT 217 a 159 c 171 g 200 t 4 others  
 ORIGIN  
 Query Match 89.2%; Score 21.4; DB 28; Length 751;  
 Best Local Similarity 91.7%; Pred. No. 1.4e+02;  
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 GTTCAGCTTTTGTACTAAGTTG 24  
 Db 633 GTTCAGCTTTTGTACTAAGTTG 610

RESULT 14  
 AQ989564/c  
 LOCUS  
 DEFINITION RfC01701 Photorhabdus luminescens strain W14 M13 library  
 Photorhabdus luminescens genomic clone PLG01701, genomic survey  
 sequence.  
 ACCESSION AQ989564  
 VERSION AQ989564.1 GI:9649458  
 KEYWORDS GSS.  
 SOURCE Photorhabdus luminescens  
 ORGANISM Photorhabdus luminescens  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Photorhabdus.  
 REFERENCE 1 (bases 1 to 672)  
 AUTHORS ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,  
 Daborn, P.J., Bowen, D. and Blattner, F.R.  
 TITLE A genomic sample sequence of the entomopathogenic bacterium  
 Photorhabdus luminescens W14: potential implications for virulence  
 JOURNAL Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)  
 MEDLINE 20378633  
 PUBMED 10919786  
 COMMENT Contact: ffrench-Constant RH  
 Department of Biology and Biochemistry  
 University of Bath  
 South Building, Bath BA2 7AY, UK  
 Tel: (44) 1225 826621  
 Fax: (44) 1225 826779  
 Email: bsarfcbath.ac.uk  
 This is one of 2,122 random reads from the M13 library. For  
 annotation of identified clones (BLASTX, BLASTN and mapping to E.  
 coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic  
 Acids Res.  
 Seq primer: M13 Forward  
 Class: shotgun.

FEATURES  
 source  
 Location/Qualifiers  
 1..672  
 /organism="Photorhabdus luminescens"  
 /mol\_type="genomic DNA"  
 /strain="W14"  
 /db\_xref="taxon:29488"  
 /clone="PLG01701"  
 /dev\_stage="primary phase variant"  
 /clone\_lib="Photorhabdus luminescens strain W14 M13  
 library"  
 /note="Genomic DNA from strain W14 was size selected (1-2  
 kb) and then cloned into M13 Janus."  
 BASE COUNT 217 a 159 c 171 g 200 t 4 others  
 ORIGIN  
 Query Match 89.2%; Score 21.4; DB 28; Length 751;  
 Best Local Similarity 91.7%; Pred. No. 1.4e+02;  
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 GTTCAGCTTTTGTACTAAGTTG 24  
 Db 633 GTTCAGCTTTTGTACTAAGTTG 610

JOURNAL RfC00126 Photorhabdus luminescens strain W14 M13 library  
 MEDLINE 20378633  
 PUBMED 10919786  
 COMMENT Contact: ffrench-Constant RH  
 Department of Biology and Biochemistry  
 University of Bath  
 South Building, Bath BA2 7AY, UK  
 Tel: (44) 1225 826621  
 Fax: (44) 1225 826779  
 Email: bsarfcbath.ac.uk  
 This is one of 2,122 random reads from the M13 library. For  
 annotation of identified clones (BLASTX, BLASTN and mapping to E.  
 coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic  
 Acids Res.  
 Seq primer: M13 Forward  
 Class: shotgun.

FEATURES  
 source  
 Location/Qualifiers  
 1..672  
 /organism="Photorhabdus luminescens"  
 /mol\_type="genomic DNA"  
 /strain="W14"  
 /db\_xref="taxon:29488"  
 /clone="PLG01701"  
 /dev\_stage="primary phase variant"  
 /clone\_lib="Photorhabdus luminescens strain W14 M13  
 library"  
 /note="Genomic DNA from strain W14 was size selected (1-2  
 kb) and then cloned into M13 Janus."  
 BASE COUNT 184 a 146 c 161 g 179 t 2 others  
 ORIGIN  
 Query Match 86.7%; Score 20.8; DB 28; Length 672;  
 Best Local Similarity 91.7%; Pred. No. 2.5e+02;  
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 GTTCAGCTTTTGTACTAAGTTG 24  
 Db 637 GTTCAGCTTTTGTACTAAGTTG 614

RESULT 15  
 AQ990861/c  
 LOCUS  
 DEFINITION RfC01698 Photorhabdus luminescens strain W14 M13 library  
 Photorhabdus luminescens genomic clone PLG01698, genomic survey  
 sequence.  
 ACCESSION AQ990861  
 VERSION AQ990861.1 GI:9649455  
 KEYWORDS GSS.  
 SOURCE Photorhabdus luminescens  
 ORGANISM Photorhabdus luminescens  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Photorhabdus.  
 REFERENCE 1 (bases 1 to 753)  
 AUTHORS ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,  
 Daborn, P.J., Bowen, D. and Blattner, F.R.  
 TITLE A genomic sample sequence of the entomopathogenic bacterium  
 Photorhabdus luminescens W14: potential implications for virulence  
 JOURNAL Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)  
 MEDLINE 20378633  
 PUBMED 10919786  
 COMMENT Contact: ffrench-Constant RH  
 Department of Biology and Biochemistry  
 University of Bath  
 South Building, Bath BA2 7AY, UK  
 Tel: (44) 1225 826621  
 Fax: (44) 1225 826779  
 Email: bsarfcbath.ac.uk  
 This is one of 2,122 random reads from the M13 library. For  
 annotation of identified clones (BLASTX, BLASTN and mapping to E.  
 coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic  
 Acids Res.  
 Seq primer: M13 Forward

Search completed: December 3, 2003, 14:49:19  
Job time : 1079.23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 652.562 Seconds  
(without alignments)  
1504.579 Million cell updates/sec

Title: US-10-082-772-4

Perfect score: 24

Sequence: 1 caacttagtatcaaaaagctgaac 24

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sv.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sv.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	24	100.0	243	6	AX092116	AX092116 Sequence
C 2	22.4	93.3	201	6	AR044609	AR044609 Sequence
C 3	22.4	93.3	201	6	E05439	E05439 Oligonucleo
C 4	22.4	93.3	201	6	I13139	I13139 Sequence 18
C 5	22.4	93.3	201	6	I36498	I36498 Sequence 18
C 6	22.4	93.3	201	6	AX092113	AX092113 Sequence
C 7	22.4	93.3	361	7	LAMINART	M23841 Bacterioph
C 8	22.4	93.3	610	6	AX101000	AX101000 Sequence
C 9	22.4	93.3	1668	9	MACHSS	D85521 Macaca fasc
C 10	22.4	93.3	2758	3	PFAHRPC	M17028 P.falciapar
C 11	22.4	93.3	2959	12	AY048723	AY048723 CRIM plas
C 12	22.4	93.3	3485	12	AF178449	AF178449 Integrati
C 13	22.4	93.3	3485	12	AF178450	AF178450 Integrati
C 14	22.4	93.3	3663	12	AY048716	AY048716 CRIM plas
C 15	22.4	93.3	3695	12	AY048733	AY048733 CRIM plas
C 16	22.4	93.3	4105	12	AF271663	AF271663 Cloning v
C 17	22.4	93.3	4190	12	XXU13848	U13848 pfxCell clo
C 18	22.4	93.3	4549	12	AF178452	AF178452 Integrati
C 19	22.4	93.3	4549	12	AF178453	AF178453 Integrati
C 20	22.4	93.3	4782	12	AY048740	AY048740 CRIM plas
C 21	22.4	93.3	5641	6	AX113748	AX113748 Sequence
C 22	22.4	93.3	5670	6	AX113749	AX113749 Sequence
C 23	22.4	93.3	5706	12	AY150267	AY150267 CRIM plas
C 24	22.4	93.3	5739	12	AY054372	AY054372 CRIM plas
C 25	22.4	93.3	5771	12	AY150262	AY150262 CRIM plas
C 26	22.4	93.3	5771	12	AY150263	AY150263 CRIM plas
C 27	22.4	93.3	5814	12	AY150265	AY150265 CRIM plas
C 28	22.4	93.3	5826	6	AX113746	AX113746 Sequence
C 29	22.4	93.3	5911	12	AY150268	AY150268 CRIM plas
C 30	22.4	93.3	5948	12	AY054373	AY054373 CRIM plas
C 31	22.4	93.3	5986	12	AY150264	AY150264 CRIM plas
C 32	22.4	93.3	6000	12	U66308	U66308 Expression
C 33	22.4	93.3	6043	6	AX370644	AX370644 Sequence
C 34	22.4	93.3	6071	6	AX113747	AX113747 Sequence
C 35	22.4	93.3	6200	12	AY150266	AY150266 CRIM plas
C 36	22.4	93.3	6664	12	AY048713	AY048713 CRIM plas
C 37	22.4	93.3	6668	12	AY048714	AY048714 CRIM plas
C 38	22.4	93.3	6742	12	AY048739	AY048739 CRIM plas
C 39	22.4	93.3	7176	12	AF178451	AF178451 Integrati
C 40	22.4	93.3	39732	7	AF069529	AF069529 Bacteriop
C 41	22.4	93.3	42529	12	CVU39284	U39284 Cloning vec
C 42	22.4	93.3	42530	12	CVU39285	U39285 Cloning vec
C 43	22.4	93.3	42531	12	CVU39286	U39286 Cloning vec
C 44	22.4	93.3	42704	12	CVU37692	U37692 Cloning vec
C 45	22.4	93.3	48502	7	LAMCG	J02459 Bacterioph

# ALIGNMENTS

RESULT 1  
AX092116/c  
LOCUS AX092116 243 bp DNA linear PAT 21-MAR-2001  
DEFINITION Sequence 5 from Patent WO0116345.  
ACCESSION AX092116  
VERSION AX092116.1 GI:13444359  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1  
AUTHORS Droegge, P.  
TITLE Sequence-specific dna recombination in eukaryotic cells  
JOURNAL Patent: WO 0116345-A 5 08-MAR-2001;  
Droegge, Peter (DE)

```
FEATURES
  source
    Location/Qualifiers
      1..243
        /organism="synthetic construct"
        /mol_type="genomic DNA"
        /db_xref="taxon:32630"
BASE COUNT      74 a      34 c      34 g      101 t
ORIGIN
Query Match      100.0%; Score 24; DB 6; Length 243;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 CAACTTAGTATCAAAAAGCTGAAC 24
      |||
Db      162 CAACTTAGTATCAAAAAGCTGAAC 139
      |||

RESULT 2
AR044609/c
LOCUS      AR044609      201 bp      DNA      linear      PAT 29-SEP-1999
DEFINITION      Sequence 18 from patent US 5817506.
ACCESSION      AR044609
VERSION      AR044609.1 GI:5966074
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unclassified.
REFERENCE      1 (bases 1 to 201)
AUTHORS      Okano,K. and Kanbara,H.
TITLE      Polynucleotide capturing support for capturing, eluting and
      collecting polynucleotides in a sample solution
JOURNAL      Patent: US 5817506-A 18 06-OCT-1998;
FEATURES
  source
    Location/Qualifiers
      1..201
        /organism="unknown"
BASE COUNT      58 a      38 c      27 g      78 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 6; Length 201;
Best Local Similarity 95.8%; Pred. No. 42;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 CAACTTAGTATCAAAAAGCTGAAC 24
      |||
Db      63 CAACTTAGTATCAAAAAGCTGAAC 40
      |||

RESULT 3
E05439/c
LOCUS      E05439      201 bp      DNA      linear      PAT 29-SEP-1997
DEFINITION      Oligonucleotide.
ACCESSION      E05439
VERSION      E05439.1 GI:2173628
KEYWORDS      JP 1993236997-A/11.
SOURCE      synthetic construct
ORGANISM      synthetic construct
      artificial sequences.
REFERENCE      1 (bases 1 to 201)
AUTHORS      Okano,K. and Kanbara,H.
TITLE      CHIP FOR CATCHING POLYNUCLEOTIDE
JOURNAL      Patent: JP 1993236997-A 11 17-SEP-1993;
      HITACHI LTD
COMMENT      OS Artificial gene
      OC Artificial sequence; Genes.
      PN JP 1993236997-A/11
      PD 17-SEP-1993
      PP 28-FEB-1992 JP 1992042829
      PI OKANO KAZUOBU, KANBARA HIDEKI
      PC C12Q1/68;
      CC strandedness: Single;
      CC topology: Linear.
FEATURES
  Location/Qualifiers
    1..243
      /organism="synthetic construct"
      /mol_type="genomic DNA"
      /db_xref="taxon:32630"
BASE COUNT      58 a      38 c      27 g      78 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 6; Length 201;
Best Local Similarity 95.8%; Pred. No. 42;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 CAACTTAGTATCAAAAAGCTGAAC 24
      |||
Db      63 CAACTTAGTATCAAAAAGCTGAAC 40
      |||

RESULT 4
I13139/c
LOCUS      I13139      201 bp      DNA      linear      PAT 26-JUL-1995
DEFINITION      Sequence 18 from patent US 5434049.
ACCESSION      I13139
VERSION      I13139.1 GI:910488
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 201)
AUTHORS      Okano,K. and Kanbara,H.
TITLE      Separation of polynucleotides using supports having a plurality of
      electrode-containing cells
JOURNAL      Patent: US 5434049-A 18 18-JUL-1995;
FEATURES
  source
    Location/Qualifiers
      1..201
        /organism="unknown"
BASE COUNT      58 a      38 c      27 g      78 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 6; Length 201;
Best Local Similarity 95.8%; Pred. No. 42;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 CAACTTAGTATCAAAAAGCTGAAC 24
      |||
Db      63 CAACTTAGTATCAAAAAGCTGAAC 40
      |||

RESULT 5
I36498/c
LOCUS      I36498      201 bp      DNA      linear      PAT 13-MAY-1997
DEFINITION      Sequence 18 from patent US 5607646.
ACCESSION      I36498
VERSION      I36498.1 GI:2086323
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 201)
AUTHORS      Okano,K. and Kanbara,H.
TITLE      Device for separating polynucleotides having a plurality of
      electrode-containing cells and movable collecting capillary
JOURNAL      Patent: US 5607646-A 18 04-MAR-1997;
FEATURES
  source
    Location/Qualifiers
      1..201
        /organism="unknown"
BASE COUNT      58 a      38 c      27 g      78 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 6; Length 201;
Best Local Similarity 95.8%; Pred. No. 42;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 CAACTTAGTATCAAAAAGCTGAAC 24
      |||
Db      63 CAACTTAGTATCAAAAAGCTGAAC 40
      |||
```

```

Db      63 CAACCTAGTATATAAAAGCTGAAC 40

RESULT 6
LOCUS      AX092113/c
DEFINITION Sequence 2 from Patent WO0116345.
ACCESSION  AX092113
VERSION     AX092113.1 GI:13444356
KEYWORDS   Bacteriophage lambda
SOURCE     Bacteriophage lambda
ORGANISM   Bacteriophage lambda
REFERENCE  Droegge,P.
AUTHORS    Droegge,P.
TITLE      Sequence-specific dna recombination in eukaryotic cells
JOURNAL    Patent: WO 0116345-A 2 08-MAR-2001;
            Droegge, Peter (DE)
FEATURES   source
            1..243
            /organism="Bacteriophage lambda"
            /mol_type="genomic DNA"
            /db_xref="taxon:10710"
BASE COUNT 74 a 34 c 33 g 102 t
ORIGIN
Query Match 93.3%; Score 22.4; DB 6; Length 243;
Best Local Similarity 95.8%; Pred. No. 41;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATATAAAAGCTGAAC 24
|||||
Db 162 CAACCTAGTATATAAAAGCTGAAC 139

RESULT 7
LOCUS      LAMINTATT
DEFINITION Bacteriophage lambda int gene region.
ACCESSION  M23841
VERSION     M23841.1 GI:215177
KEYWORDS   site-specific recombination.
SOURCE     Bacteriophage lambda
ORGANISM   Bacteriophage lambda
REFERENCE  Viruses: dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
            Lambda-like viruses.
            1 (bases 1 to 361)
            Schmeissner,U., McKenney,K., Rosenberg,M. and Court,D.
            Removal of a terminator structure by RNA processing regulates int
            gene expression
            J. Mol. Biol. 176 (1), 39-53 (1984)
JOURNAL    MEDLINE
MEDLINE    84242838
PUBMED     6234400
COMMENT    Original source text: Bacteriophage lambda DNA.
FEATURES   Location/Qualifiers
            1..361
            /organism="Bacteriophage lambda"
            /mol_type="genomic DNA"
            /db_xref="taxon:10710"
BASE COUNT 134 a 60 c 60 g 107 t
ORIGIN
Query Match 93.3%; Score 22.4; DB 7; Length 361;
Best Local Similarity 95.8%; Pred. No. 39;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATATAAAAGCTGAAC 24
|||||
Db 98 CAACCTAGTATATAAAAGCTGAAC 121

RESULT 8

```

```

AX101000/c
LOCUS      AX101000
DEFINITION Sequence 1 from Patent WO0121780.
ACCESSION  AX101000
VERSION     AX101000.1 GI:13619857
KEYWORDS   Nicotiana tabacum (common tobacco)
SOURCE     Nicotiana tabacum
ORGANISM   Nicotiana tabacum
REFERENCE  Meyer,P.L. and Zubko,E.L.
AUTHORS    Meyer,P.L. and Zubko,E.L.
TITLE      Targeted gene removal
JOURNAL    Patent: WO 0121780-A 1 29-MAR-2001;
            THE UNIVERSITY OF LEEDS (GB)
FEATURES   Location/Qualifiers
            1..610
            /organism="Nicotiana tabacum"
            /mol_type="genomic DNA"
            /db_xref="taxon:4097"
BASE COUNT 164 a 108 c 110 g 228 t
ORIGIN
Query Match 93.3%; Score 22.4; DB 6; Length 610;
Best Local Similarity 95.8%; Pred. No. 35;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATATAAAAGCTGAAC 24
|||||
Db 300 CAACCTAGTATATAAAAGCTGAAC 277

RESULT 9
LOCUS      MACHSS
DEFINITION Macaca fascicularis mRNA for hydroxysteroid sulfotransferase
            subunit, complete cds.
ACCESSION  D85521
VERSION     D85521.1 GI:1345405
KEYWORDS   hydroxysteroid sulfotransferase subunit.
SOURCE     Macaca fascicularis (Crab-eating macaque)
ORGANISM   Macaca fascicularis
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopitheinae; Macaca.
            1 (bases 1 to 1668)
            Ogura,K., Satsukawa,M., Kato,K., Okuda,H. and Watabe,T.
            Molecular cloning of monkey liver hydroxysteroid sulfotransferase
            Unpublished
            2 (bases 1 to 1668)
            Ogura,K.
            Direct Submission
            Submitted (23-MAY-1996) Kenichiro Ogura, Tokyo University of
            Pharmacy and Life Science, Department of Drug Metabolism and
            Molecular Toxicology, 1432-1 Horinouchi, Hachioji, Tokyo 192-0392,
            Japan (E-mail:ogurak@ps.toyaku.ac.jp, Tel.:+81-426-76-4518,
            Fax: +81-426-76-4517)
            Location/Qualifiers
            1..1668
            /organism="Macaca fascicularis"
            /mol_type="mRNA"
            /db_xref="taxon:9541"
            /clone="monHST-1"
            /sex="male"
            /tissue_type="liver"
            /clone_lib="lambda gt11"
            611..1468
            /codon_start=1
            /product="hydroxysteroid sulfotransferase subunit"
            /protein_id="BAA12823.1"
            /db_xref="GI:1345406"
            /translation="MSDDLFWFEGIAFPNMFSETRLRKVRDFVKDEIVLTVPK"
CDS

```

```

SGTNWLEILCLHNSGDPKWKIOSVPIWERSPMWETEMGYKLLSBEQGRPLFSHLPI
QLPKPFFSKAKVILMENPRDFVUSGYFPMNSUKFVKPKSMQOYEMFCQGVIVY
GSWFDHGWMPREKKNFLLSYELKQDCTERTVEKICQFKGKLEFELNLIKNS
SFSQMKENKNSLSDVDFVEKAQLLRKISGDWKNHLLTVAQAFAPKLFQEKWTD
LPRELFPWE"
BASE COUNT      521 a   329 c   359 g   459 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 9; Length 1668;
Best Local Similarity 95.8%; Pred. No. 30;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATCAAAAAGCTGAAC 24
|||||
Db 230 CAACCTAGTATCAAAAAGCTGAAC 253

RESULT 10
PFAHRPC
LOCUS PFAHRPC 2758 bp DNA linear INV 26-APR-1993
DEFINITION P.falciparum histidine-rich protein genes.
ACCESSION MI7028
VERSION MI7028.1 GI:160339
KEYWORDS histidine-rich protein.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 2758)
AUTHORS Lenstra,R., d'Auriol,L., Andrieu,B., Le Bras,J. and Galibert,F.
TITLE Cloning and sequencing of Plasmodium falciparum DNA fragments
containing repetitive regions potentially coding for histidine-rich
proteins: identification of two overlapping reading frames
JOURNAL Biochem. Biophys. Res. Commun. 146 (1), 368-377 (1987)
MEDLINE 87270765
PUBMED 3038111
COMMENT Original source text: P.falciparum, cDNA to mRNA, clone M4.
FEATURES
Location/Qualifiers
1..2758
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
1..234
/organism="histidine rich protein A"
<1..234
/organism="histidine rich protein A"
/codon_start=1
/product="histidine rich protein A"
/protein_id="AAA29618.1"
/db_xref="GI:160340"
/translation="DPYKERIKSDIRQINESQYKSLAYKYISGEDYTOYLLNEVLK
DDQYCTCTRTTVEESMDNTVEFAKKQVELSA"
824..1072
/organism="histidine rich protein B"
824..1072
/organism="histidine rich protein B"
/codon_start=1
/product="histidine rich protein B"
/protein_id="AAA29619.1"
/db_xref="GI:160341"
/translation="MLNHRVHYFRRHHHNLNHLHYRHHHHRRHHRRHHRRHHQILHQN
RHOIHOILLNNKIMGYAIFLFLSILLHLYVLIHRL"
complement(828..1589)
/organism="histidine rich protein E"
complement(828..1589)
/organism="histidine rich protein E"
/codon_start=1
/product="histidine rich protein E"
/protein_id="AAA29620.1"
/db_xref="GI:160342"
/translation="WYLFYIYFFFFVIVQKDIQIDIKCAEQNNIQKQYDE
KKKPLFEKRDEIIQIPGWANTLRKHPALSDIVPEDIDILNHLVDLKKONMNNGS
YKTIPIGEKAKEFELTLVKHVTFDNNQEKVECTRIRKWKSGKNPIAAVTNRSDL
DNEIPKWSIFEFWFTDELQDKPDVGELIRREIWHNPLSYGLGEFDFDDDFDEBFD

```

```

DDDDDDDDDDDDDDKDDDLDCGDDGNNDDND"
complement(2141..2650)
/organism="Bacteriophage lambda"
/notes="putative VECTOR sequence Bacteriophage lambda
(J02459); putative"
2525..2758
/organism="histidine rich protein C"
2525..2758
/organism="histidine rich protein C"
/codon_start=1
/product="histidine rich protein C"
/protein_id="AAA29621.1"
/db_xref="GI:160343"
/translation="MVLVTGNRALAQDFCLLALIFCHQTCRTPEKHAKSQSSAKLV
INISLIISHRLHPRRRQHHNNFAPTWYWG"
complement(2596..2758)
/organism="histidine rich protein D"
complement(2596..2758)
/organism="histidine rich protein D"
/codon_start=2
/product="histidine rich protein D"
/protein_id="AAA29622.1"
/db_xref="GI:160344"
/translation="PQYFVGAKLFRWCRWRRRRWLVKMLLIETSFALDCEA
LCFSGVRQV"
BASE COUNT      1004 a   404 c   311 g   1039 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 3; Length 2758;
Best Local Similarity 95.8%; Pred. No. 28;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATCAAAAAGCTGAAC 24
|||||
Db 2387 CAACCTAGTATCAAAAAGCTGAAC 2410

RESULT 11
AY048723/c
LOCUS AY048723 2959 bp DNA circular SYN 17-OCT-2001
DEFINITION CRIM plasmid pAH120, complete sequence.
ACCESSION AY048723
VERSION AY048723.1 GI:16209102
KEYWORDS CRIM plasmid pAH120
SOURCE CRIM plasmid pAH120
ORGANISM CRIM plasmid pAH120
REFERENCE 1 (bases 1 to 2959)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Conditional-replication, integration, excision, and retrieval
plasmid-host systems for gene structure-function studies of
bacteria
JOURNAL J. Bacteriol. 183 (21), 6384-6393 (2001)
MEDLINE 21475783
PUBMED 11591683
REFERENCE 2 (bases 1 to 2959)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Direct Submision
JOURNAL Submitted (29-JUN-2001) Biological Sciences, Purdue University,
Lilly Hall, West Lafayette, IN 47907, USA
FEATURES
Location/Qualifiers
1..2959
/organism="CRIM plasmid pAH120"
/mol_type="genomic DNA"
/db_xref="taxon:172348"
complement(33..320)
/organism="Prhab"
complement(1719..2513)
/codon_start=1
/transl_table=11
/product="kan"
/protein_id="AAL09869.1"
/db_xref="GI:16209103"
/translation="MIEQDLHAGSPAAWVERLFGYDWAQQTICSDAAVFLSAQGR

```

PVLVFKTDSGALNELODEAARLSWLATTGVCFAVLDDVTAGRDWILLGVEPGQDL  
 LSSHIAPEAKVYSIMADARRLHTLDPATCPDHOAKHRIERATRMAGLVDQDLDE  
 EHQGLAPRELARLKARMPDGEDLVVTHGDACLINWVENGRFSGFDICGLGVADRY  
 QDIALADIAEELGGEWADRFLVYIGIAADPSQRIAFYRLDDEFF"  
 BASE COUNT 767 a 718 c 716 g 758 t  
 ORIGIN

Query Match 93.3%; Score 22.4; DB 12; Length 2959;  
 Best Local Similarity 95.8%; Pred. No. 28;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
 |||||  
 Db 1102 CAACCTAGTATCAAAAAGCTGAAC 1079

RESULT 12  
 AFI178449/c  
 LOCUS AFI178449 3485 bp DNA linear SYN 20-JUL-2000  
 DEFINITION Integration vector pCD11PKS chloramphenicol transacetylase (cat)  
 and beta-galactosidase alpha peptide (lacZa) genes, complete cds.

ACCESSION AFI178449  
 VERSION AFI178449.1 GI:9294785

KEYWORDS Integration vector pCD11PKS  
 SOURCE Integration vector pCD11PKS  
 ORGANISM Artificial sequences; vectors.

REFERENCE 1 (bases 1 to 3485)

AUTHORS Platt,R., Drescher,C., Park,S.K. and Phillips,G.J.

TITLE Genetic system for reversible integration of DNA constructs and

lacZ gene fusions into the Escherichia coli chromosome

JOURNAL Plasmid 43 (1), 12-23 (2000)

MEDLINE 20079288

PUBMED 10610816

REFERENCE 2 (bases 1 to 3485)

AUTHORS Platt,R., Drescher,C., Park,S.K. and Phillips,G.J.

TITLE Direct Submission

JOURNAL Submitted (18-AUG-1999) Microbiology, Iowa State University, 207

Science I Building, Ames, IA 50011, USA

FEATURES Location/Qualifiers

1..3485

/organism="Integration vector pCD11PKS"

/mol\_type="genomic DNA"

/db\_xref="taxon:106602"

/lab\_host="Escherichia coli"

1..385

/note="R6Kgamma"

688..915

/note="attP; attachment site from bacteriophage lambda"

complement(1327..1986)

/gene="cat"

complement(1327..1986)

/gene="cat"

/codon\_start=1

/transl\_table=11

/product="chloramphenicol transacetylase"

/protein\_id="AAF86671.1"

/db\_xref="GI:9294787"

/translation="MEKKITGYTVTDISQWHRKEHFEAFQSVQAQCTYNQTVQDLITAF

LKTVKNKHKFPATFIHLARLNARHPEFRMAKDGELVWDSVHPCYTVFHEQTET

SLWSYHDDFRQFLHIYSQDVACGENLAYFPKGIENMFVSANPWVSTFDLNV

ANNDNFPAPVFTMGKYTQGDVKVLMPLAIQVHVAVCDGFHVGRLNELQYCDQVQGG

A"

complement(2754..3137)

/gene="lacZa"

complement(2754..3137)

/gene="lacZa"

/codon\_start=1

/transl\_table=11

/product="beta-galactosidase alpha peptide"

/protein\_id="AAF86670.1"

/db\_xref="GI:9294786"

/translation="MTMITPSAQLTITKGNKSWVGPSPRSNVSVLSISNCSPODPL

VLERPPRWSSNPYSSEYARSALVAVLQRRDWNPGVTQLNRLAAHPFPASRWNSSE  
 ARTDRSQQLRSINGEWLFRPVAH"  
 2972..3081

/note="multiple cloning site"  
 BASE COUNT 968 a 797 c 803 g 917 t  
 ORIGIN

Query Match 93.3%; Score 22.4; DB 12; Length 3485;  
 Best Local Similarity 95.8%; Pred. No. 27;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
 |||||  
 Db 834 CAACCTAGTATCAAAAAGCTGAAC 811

RESULT 13  
 AFI178450/c  
 LOCUS AFI178450 3485 bp DNA linear SYN 20-JUL-2000  
 DEFINITION Integration vector pCD11PSK chloramphenicol transacetylase (cat)  
 and beta-galactosidase alpha peptide (lacZa) genes, complete cds.

ACCESSION AFI178450  
 VERSION AFI178450.1 GI:9294788

KEYWORDS Integration vector pCD11PSK  
 SOURCE Integration vector pCD11PSK  
 ORGANISM Artificial sequences; vectors.

REFERENCE 1 (bases 1 to 3485)

AUTHORS Platt,R., Drescher,C., Park,S.K. and Phillips,G.J.

TITLE Genetic system for reversible integration of DNA constructs and

lacZ gene fusions into the Escherichia coli chromosome

JOURNAL Plasmid 43 (1), 12-23 (2000)

MEDLINE 20079288

PUBMED 10610816

REFERENCE 2 (bases 1 to 3485)

AUTHORS Platt,R., Drescher,C., Park,S.K. and Phillips,G.J.

TITLE Direct Submission

JOURNAL Submitted (18-AUG-1999) Microbiology, Iowa State University, 207

Science I Building, Ames, IA 50011, USA

FEATURES Location/Qualifiers

1..3485

/organism="Integration vector pCD11PSK"

/mol\_type="genomic DNA"

/db\_xref="taxon:106603"

/lab\_host="Escherichia coli"

1..385

/note="R6Kgamma"

688..915

/note="attP; attachment site from bacteriophage lambda"

complement(1327..1986)

/gene="cat"

complement(1327..1986)

/gene="cat"

/codon\_start=1

/transl\_table=11

/product="chloramphenicol transacetylase"

/protein\_id="AAF86673.1"

/db\_xref="GI:9294790"

/translation="MEKKITGYTVTDISQWHRKEHFEAFQSVQAQCTYNQTVQDLITAF

LKTVKNKHKFPATFIHLARLNARHPEFRMAKDGELVWDSVHPCYTVFHEQTET

SLWSYHDDFRQFLHIYSQDVACGENLAYFPKGIENMFVSANPWVSTFDLNV

ANNDNFPAPVFTMGKYTQGDVKVLMPLAIQVHVAVCDGFHVGRLNELQYCDQVQGG

A"

complement(2754..3137)

/gene="lacZa"

complement(2754..3137)

/gene="lacZa"

/codon\_start=1

/transl\_table=11

/product="beta-galactosidase alpha peptide"

/protein\_id="AAF86672.1"

/db\_xref="GI:9294789"

/translation="MTMITPSAQLTITKGNKSWSTAVAAALELVDPGGRNSISLS

```
IPSTRSGGVPNSPVSYARSIAVLVLRQDWDENPGVTQNLRLAHPFPASWNSSE
ARTDRPSQOLSLNGELWLRPVAH"
2972..3081
/misc_feature 967 a 798 c 802 g 918 t
BASE COUNT 967 a 798 c 802 g 918 t
ORIGIN
Query Match 93.3%; Score 22.4; DB 12; Length 3485;
Best Local Similarity 95.8%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACTAGTATCAAAAAGCTGAAC 24
|||||
Db 834 CAACTAGTATCAAAAAGCTGAAC 811
|||||

RESULT 14
AY048716/c 3663 bp DNA circular SYN 17-OCT-2001
LOCUS
DEFINITION CRIM plasmid pAH63, complete sequence.
ACCESSION AY048716
VERSION AY048716.1 GI:16209075
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 3663)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Conditional-replication, integration, excision, and retrieval
of plasmid-host systems for gene structure-function studies of
bacteria
JOURNAL J. Bacteriol. 183 (21), 6384-6393 (2001)
MEDLINE 21475783
PUBMED 11591683
REFERENCE 2 (bases 1 to 3663)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-2001) Biological Sciences, Purdue University,
Lilly Hall, West Lafayette, IN 47907, USA
FEATURES
Location/Qualifiers
1..3663
/organism="CRIM plasmid pAH63"
/mol_type="genomic DNA"
/db_xref="taxon:172341"
101..1030
/gene="pstS"
101..1030
/gene="pstS"
/codon_start=1
/transl_table=11
/product="pstS"
/protein_id="AAL09852.1"
/db_xref="GI:16209076"
/translation="MKMRTVATVVAATLSMSAFVFAEASLTGAGTTFAPVAKW
ADYQKRTGNVYGGSGGKVIANTVDFGASDAPLSDEKLAGLGFQFTVIG
GVVLAVNIPGLKSELVGLDGLVIGIKWDDDEAIKLNPLKPSQIAVVR
ADGSGTFFVTSYLAKEWKNVGTSTVKWPIGLGKGNDAFAFVQRLPGAIGY
VEYAKONNLAVTKLISADGKPVSTPEENFANAAGADWSKTFAQDLINQGEDAMP
ITSTTFLIHQDKPEQDSAILRLGVQNRG"
complement(2423..3217)
/codon_start=1
/transl_table=11
/product="kan"
/protein_id="AAL09853.1"
/db_xref="GI:16209077"
/translation="MIEQDLHAGSPAANVERLFGYDWAQQTIGCSDAAVFLSAQGR
PVLVKTDLSGALNEIQDEARLSWLTGVPAAVLDVVTEAGRDWLLIGVPGDGL
LSSLAPAEKVSIMADARLRLTDPATCFDQAKHRIERARTMEAGLVDDDLDE
EHQGLAPAEFLKARMPDGEDLVVTHGDACLPLNIMVENGFSGFIDCGRLGVADRY
QDIALATRIAELGGEWADRFLVLYGIAAPDSQRIAFYLLDEFF"
complement(3174..3663)
/notes="PstynI; from pBH7013"
BASE COUNT 954 a 892 c 914 g 903 t

misc_feature 967 a 798 c 802 g 918 t
BASE COUNT 967 a 798 c 802 g 918 t
ORIGIN
Query Match 93.3%; Score 22.4; DB 12; Length 3485;
Best Local Similarity 95.8%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACTAGTATCAAAAAGCTGAAC 24
|||||
Db 1806 CAACTAGTATCAAAAAGCTGAAC 1783
|||||

RESULT 15
AY048733/c 3695 bp DNA circular SYN 17-OCT-2001
LOCUS
DEFINITION CRIM plasmid pAH150, complete sequence.
ACCESSION AY048733
VERSION AY048733.1 GI:16209150
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 3695)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Conditional-replication, integration, excision, and retrieval
of plasmid-host systems for gene structure-function studies of
bacteria
JOURNAL J. Bacteriol. 183 (21), 6384-6393 (2001)
MEDLINE 21475783
PUBMED 11591683
REFERENCE 2 (bases 1 to 3695)
AUTHORS Haldemann,A. and Wanner,B.L.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-2001) Biological Sciences, Purdue University,
Lilly Hall, West Lafayette, IN 47907, USA
FEATURES
Location/Qualifiers
1..3695
/organism="CRIM plasmid pAH150"
/mol_type="genomic DNA"
/db_xref="taxon:172358"
complement(38..193)
/notes="arac' terminal fragment"
198..507
/notes="arabB3p"
548..1237
/codon_start=1
/transl_table=11
/product="Phob"
/protein_id="AAL09901.1"
/db_xref="GI:16209152"
/translation="MARRILVDEADPIREWCVELEQNGFQVFEADYDSAVNQLNE
PWPDLILLDMWPGSGIQIKHLKRESMTDIPVWMLTARGEEDRVRLGETGADY
ITKFPSPKELVARIKAVMRRIISPMAVEEIVEMQGLSDPTSHRVMAGEEPLMGPTF
KLLFFHTHPERVYSREQLLNHVMTGNTVYVEDRTVDVHIRRLKALEPGGHRMVQTV
RGTYRFRSTRF"
complement(2455..3249)
/codon_start=1
/transl_table=11
/product="kan"
/protein_id="AAL09900.1"
/db_xref="GI:16209151"
/translation="MIEQDLHAGSPAANVERLFGYDWAQQTIGCSDAAVFLSAQGR
PVLVKTDLSGALNEIQDEARLSWLTGVPAAVLDVVTEAGRDWLLIGVPGDGL
LSSLAPAEKVSIMADARLRLTDPATCFDQAKHRIERARTMEAGLVDDDLDE
EHQGLAPAEFLKARMPDGEDLVVTHGDACLPLNIMVENGFSGFIDCGRLGVADRY
QDIALATRIAELGGEWADRFLVLYGIAAPDSQRIAFYLLDEFF"
complement(914 c 913 g 919 t
BASE COUNT 949 a 914 c 913 g 919 t
ORIGIN
Query Match 93.3%; Score 22.4; DB 12; Length 3695;
Best Local Similarity 95.8%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACTAGTATCAAAAAGCTGAAC 24
```



Db 1838 CAACTTAGTATATAAAAAAGCTGAAC 1815  
|||||

Search completed: December 3, 2003, 13:18:43  
Job time : 653.562 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 28.562 Seconds  
(without alignments)  
370.884 Million cell updates/sec

Title: US-10-082-772-4

Perfect score: 24  
Sequence: 1 caacttagtatcaaaagctgaac 24

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22.4	93.3	201	1	US-08-021-667A-18
C 2	22.4	93.3	201	1	US-08-410-544-18
C 3	22.4	93.3	201	1	US-08-728-785A-18
C 4	22.4	93.3	4909	3	US-08-556-978B-18
C 5	22.4	93.3	7652	1	US-07-590-988A-1
C 6	21.2	88.3	25	3	US-09-296-280-43
C 7	20.4	85.0	25	3	US-09-296-280-42
C 8	19.4	80.8	21	3	US-08-894-511-7
C 9	19.4	80.8	21	4	US-09-655-728-7
C 10	19.2	80.0	25	3	US-09-233-433-15
C 11	19.2	80.0	25	3	US-09-005-476-15
C 12	19.2	80.0	25	3	US-09-233-492-15
C 13	19.2	80.0	25	3	US-09-296-280-15
C 14	19.2	80.0	25	4	US-09-438-074-15
C 15	19.2	80.0	25	5	PCT-US96-10082A-15
C 16	18.6	77.5	25	3	US-09-233-493-5
C 17	18.6	77.5	25	3	US-09-005-476-5
C 18	18.6	77.5	25	3	US-09-233-492-5
C 19	18.6	77.5	25	3	US-09-296-280-5
C 20	18.6	77.5	25	4	US-09-438-074-5
C 21	18.6	77.5	25	5	PCT-US96-10082A-5
C 22	18.4	76.7	25	3	US-09-296-280-39
C 23	18.2	75.8	580073	4	US-08-545-528D-1
C 24	17.8	74.2	25	3	US-09-296-280-41
C 25	17.8	74.2	228	4	US-09-107-532A-667
C 26	17.8	74.2	1123	3	US-09-188-930-28
C 27	17.8	74.2	1123	3	US-09-188-930-203

C 28	17.8	74.2	1123	4	US-09-312-283C-28	Sequence 28, Appl
C 29	17.8	74.2	1123	4	US-09-312-283C-203	Sequence 203, Appl
C 30	17.8	74.2	2408	1	US-08-486-013-69	Sequence 69, Appl
C 31	17.8	74.2	2408	2	US-08-482-279-69	Sequence 69, Appl
C 32	17.8	74.2	2408	2	US-08-342-268-69	Sequence 69, Appl
C 33	17.8	74.2	2408	3	US-09-015-968-69	Sequence 69, Appl
C 34	17.8	74.2	2408	4	US-09-397-386-69	Sequence 69, Appl
C 35	17.8	74.2	3484	4	US-09-380-090A-1	Sequence 1, Appl
C 36	17.8	74.2	3484	4	US-09-016-366A-13	Sequence 13, Appl
C 37	17.8	74.2	3757	2	US-08-978-404B-19	Sequence 19, Appl
C 38	17.8	74.2	5349	4	US-09-068-101-7	Sequence 7, Appl
C 39	17.8	74.2	5611	4	US-09-068-101-10	Sequence 10, Appl
C 40	17.6	73.3	25	3	US-09-233-493-9	Sequence 9, Appl
C 41	17.6	73.3	25	3	US-09-233-493-11	Sequence 11, Appl
C 42	17.6	73.3	25	3	US-09-233-493-16	Sequence 16, Appl
C 43	17.6	73.3	25	3	US-09-005-476-9	Sequence 9, Appl
C 44	17.6	73.3	25	3	US-09-005-476-11	Sequence 11, Appl
C 45	17.6	73.3	25	3		

ALIGNMENTS

RESULT 1  
US-08-021-667A-18/c  
; Sequence 18, Application US/08021667A  
; Patent No. 5434049  
; GENERAL INFORMATION:  
; APPLICANT: Okano, Kazunori  
; APPLICANT: Kambata, Hideki  
; TITLE OF INVENTION: POLYNUCLEOTIDE CAPTURING TIP AND  
; TITLE OF INVENTION: POLYNUCLEOTIDE PREPARATIVE METHOD AND DETECTION  
; TITLE OF INVENTION: METHOD USING SAME  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Antonelli, Terry, Stout & Kraus  
; STREET: Suite 600, 1919 Pennsylvania Ave., NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/021,667A  
; FILING DATE: 19930224  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Terry, David T.  
; REGISTRATION NUMBER: 20,178  
; REFERENCE/DOCKET NUMBER: 520.31930X00  
; TELEPHONE: 202-828-0300  
; TELEFAX: 202-828-0380  
; TELEX: 440380/248545  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 201 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: YES  
; ANTI-SENSE: NO  
US-08-021-667A-18

Query Match 93.3%; Score 22.4; DB 1; Length 201;  
Best Local Similarity 95.8%; Pred. No. 0.28;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
|||||  
Db 63 CAACCTTAGTATCAAAAAGCTGAAC 40

## RESULT 2

US-08-410-544-18/c  
; Sequence 18, Application US/08410544  
; Patent No. 5607646  
; GENERAL INFORMATION:  
; APPLICANT: Okano, Kazunori  
; APPLICANT: Kambara, Hideki  
; TITLE OF INVENTION: POLYNUCLEOTIDE CAPTURING TIP AND  
; TITLE OF INVENTION: POLYNUCLEOTIDE PREPARATIVE METHOD AND DETECTION  
; TITLE OF INVENTION: METHOD USING SAME  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Antonelli, Terry, Stout & Kraus  
; STREET: Suite 600, 1919 Pennsylvania Ave., NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/410,544  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/021,667  
; FILING DATE: 24-FEB-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Terry, David T.  
; REGISTRATION NUMBER: 20,178  
; REFERENCE/DOCKET NUMBER: 520.31930X00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-828-0300  
; TELEFAX: 202-828-0380  
; TELEX: 248545  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 201 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEetical: YES  
; ANTI-SENSE: NO  
US-08-410-544-18

Query Match 93.3%; Score 22.4; DB 1; Length 201;  
Best Local Similarity 95.8%; Pred. No. 0.28;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
|||||  
Db 63 CAACCTTAGTATCAAAAAGCTGAAC 40

## RESULT 3

US-08-728-785A-18/c  
; Sequence 18, Application US/08728785A  
; Patent No. 5817506  
; GENERAL INFORMATION:  
; APPLICANT: Okano, Kazunori  
; APPLICANT: Kambara, Hideki  
; TITLE OF INVENTION: POLYNUCLEOTIDE CAPTURING TIP AND  
; TITLE OF INVENTION: POLYNUCLEOTIDE PREPARATIVE METHOD AND DETECTION  
; TITLE OF INVENTION: METHOD USING SAME

NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Antonelli, Terry, Stout & Kraus  
STREET: Suite 1800, 1300 No. 5817506th Seventeenth St.  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22209  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/728,785A  
FILING DATE: 10-OCT-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/410,544  
FILING DATE: 21-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/021,667  
FILING DATE: 24-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Terry, David T.  
REGISTRATION NUMBER: 20,178  
REFERENCE/DOCKET NUMBER: 520.31930X00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-312-6660  
TELEFAX: 703-312-6666  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHEtical: YES  
ANTI-SENSE: NO  
US-08-728-785A-18

Query Match 93.3%; Score 22.4; DB 1; Length 201;  
Best Local Similarity 95.8%; Pred. No. 0.28;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
|||||  
Db 63 CAACCTTAGTATCAAAAAGCTGAAC 40

## RESULT 4

US-08-556-978B-78  
; Sequence 78, Application US/08556978B  
; Patent No. 6268169  
; GENERAL INFORMATION:  
; APPLICANT: FAHNESTOCK, STEPHEN F.  
; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED  
; TITLE OF INVENTION: SPIDER SILK ANALOGS  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
; STREET: 1007 MARKET STREET  
; CITY: WILMINTON  
; STATE: DELAWARE  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 19898  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.50 INCH  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: MICROSOFT WINDOWS 95  
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/556,978B

;; FILING DATE: 315  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/077,600  
;; FILING DATE: JUNE 15, 1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: FLOYD, LINDA AXAMETHY  
;; REGISTRATION NUMBER: 33,692  
;; REFERENCE/DOCKET NUMBER: CR-9389-A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 302-892-8112  
;; TELEFAX: 302-773-0164  
;; INFORMATION FOR SEQ ID NO: 78:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4909 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
US-08-556-978B-78  
  
Query Match 93.3%; Score 22.4; DB 3; Length 4909;  
Best Local Similarity 95.8%; Pred. No. 0.32;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
Db 2355 CAACCTAGTATCAAAAAGCTGAAC 2378  
  
RESULT 5  
US-07-590-988A-1/c  
; Sequence 1, Application US/07590988A  
; Patent No. 5227288  
; GENERAL INFORMATION:  
; APPLICANT: Blattner, Frederick R.  
; TITLE OF INVENTION: DNA Sequencing Vector with  
; TITLE OF INVENTION: Reversible Insert  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles and Brady  
; STREET: P.O. BOX 2113  
; STREET: FIRST WISCONSIN PLAZA  
; CITY: MADISON  
; STATE: WISCONSIN  
; COUNTRY: U.S.A.  
; ZIP: 53701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette - 3.50 inch, 800Kb storage  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh  
; SOFTWARE: Microsoft Word 4.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07590,988A  
; FILING DATE: 19901001  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Nicholas J. Seay  
; REGISTRATION NUMBER: 27,385  
; REFERENCE/DOCKET NUMBER: 9629691682  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (608) 251-5000  
; TELEFAX: (608) 251-9166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7652 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: Other nucleic acid;  
; DESCRIPTION: synthetic recombinant plasmid  
; HYPOTHEICAL: no  
; ANTI-SENSE: no

US-07-590-988A-1  
  
Query Match 93.3%; Score 22.4; DB 1; Length 7652;  
Best Local Similarity 95.8%; Pred. No. 0.33;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
Db 6094 CAACCTAGTATCAAAAAGCTGAAC 6071  
  
RESULT 6  
US-09-296-280-43/c  
; Sequence 43, Application US/09296280  
; Patent No. 6277608  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; APPLICANT: Temple, Gary F.  
; APPLICANT: Fox, Donna K.  
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having  
; TITLE OF INVENTION: Recombination Sites  
; FILE REFERENCE: 0942.2850007  
; CURRENT APPLICATION NUMBER: US/09/296,280  
; CURRENT FILING DATE: 1999-04-22  
; EARLIER APPLICATION NUMBER: US 09/177,387  
; EARLIER FILING DATE: 1998-10-23  
; EARLIER APPLICATION NUMBER: US 60/065,930  
; EARLIER FILING DATE: 1997-10-24  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 43  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: recombination  
; OTHER INFORMATION: products  
US-09-296-280-43  
  
Query Match 88.3%; Score 21.2; DB 3; Length 25;  
Best Local Similarity 83.3%; Pred. No. 0.87;  
Matches 20; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
Db 24 CAACCTAGTATCAAAAAGCTGAAC 1  
  
RESULT 7  
US-09-296-280-42/c  
; Sequence 42, Application US/09296280  
; Patent No. 6277608  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; APPLICANT: Temple, Gary F.  
; APPLICANT: Fox, Donna K.  
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having  
; TITLE OF INVENTION: Recombination Sites  
; FILE REFERENCE: 0942.2850007  
; CURRENT APPLICATION NUMBER: US/09/296,280  
; CURRENT FILING DATE: 1999-04-22  
; EARLIER APPLICATION NUMBER: US 09/177,387  
; EARLIER FILING DATE: 1998-10-23  
; EARLIER APPLICATION NUMBER: US 60/065,930  
; EARLIER FILING DATE: 1997-10-24  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 42  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Unknown

```
/
/
/ OTHER INFORMATION: Description of Unknown Organism: recombination
/
/ OTHER INFORMATION: products
US-09-296-280-42

Query Match      85.0%; Score 20.4; DB 3; Length 25;
Best Local Similarity 75.0%; Pred. No. 1.9;
Matches 18; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATCAAAAGCTGAC 24
   ||:||||:||||:||||:||||:
Db 24 CMASTTGTATATAAAAGCTGAC 1

RESULT 8
US-08-894-511-7/c
; Sequence 7, Application US/08894511
; Patent No. 6143530
; GENERAL INFORMATION:
; APPLICANT: CROUZET, Joel
; APPLICANT: SCHERMAN, Daniel
; APPLICANT: CAMERON, Beatrice
; APPLICANT: WILS, Pierre
; APPLICANT: DARQUET, Anne-Marie
; TITLE OF INVENTION: DNA MOLECULES, PREPARATION THEREOF AND
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,511
; FILING DATE: 23-FEB-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02117
; FILING DATE: 23-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR96/00274
; FILING DATE: 21-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky Esq., Martin F.
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: ST95013-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3816
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Oligonucleotide"
US-08-894-511-7

Query Match      80.8%; Score 19.4; DB 3; Length 21;
Best Local Similarity 95.2%; Pred. No. 5.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATCAAAAGCTG 21
   |||||
Db 21 CAACCTAGTATCAAAAGCTG 1

RESULT 9
US-09-655-728-7/c
; Sequence 7, Application US/09655728
; Patent No. 6492164
; GENERAL INFORMATION:
; APPLICANT: CROUZET, Joel
; APPLICANT: SCHERMAN, Daniel
; APPLICANT: CAMERON, Beatrice
; APPLICANT: WILS, Pierre
; APPLICANT: DARQUET, Anne-Marie
; TITLE OF INVENTION: DNA MOLECULES, PREPARATION THEREOF AND
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/655,728
; FILING DATE: 05-Sep-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/894,511
; FILING DATE: <Unknown>
; APPLICATION NUMBER: WO FR96/00274
; FILING DATE: 21-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky Esq., Martin F.
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: ST95013-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3816
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Oligonucleotide"
US-09-655-728-7

Query Match      80.8%; Score 19.4; DB 4; Length 21;
Best Local Similarity 95.2%; Pred. No. 5.3;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATCAAAAGCTG 21
   |||||
Db 21 CAACCTAGTATCAAAAGCTG 1

RESULT 10
US-09-233-493-15/c
; Sequence 15, Application US/09233493
; Patent No. 6143557
; GENERAL INFORMATION:
; APPLICANT: Hartley, James L.
; APPLICANT: Brasch, Michael A.
; TITLE OF INVENTION: Recombinational Cloning Using Engineered
; TITLE OF INVENTION: Recombination Sites
; NUMBER OF SEQUENCES: 35
```

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C  
;; STREET: 1100 New York Ave., N. W. Suite 600  
;; CITY: Washington  
;; STATE: DC  
;; COUNTRY: USA  
;; ZIP: 20005-3934  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/233,493  
;; FILING DATE: 20-JAN-1999  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 09/005,476  
;; FILING DATE: 12-JAN-1998  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/486,139  
;; FILING DATE: 07-JUN-1995  
;; CLASSIFICATION:  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-2600  
;; TELEFAX: 202-371-2540  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 25 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: both  
;; TOPOLOGY: both  
;; MOLECULE TYPE: cdna  
US-09-233-493-15

Query Match 80.0%; Score 19.2; DB 3; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CAACTTAGTATCAAAAAGCTGAAC 24  
|||||  
Db 24 CAACTTTGTACAAAAAAGCTGAAC 1

RESULT 11  
US-09-005-476-15/c  
; Sequence 15, Application US/09005476  
; Patent No. 6171861  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; TITLE OF INVENTION: Recombinational Cloning Using Engineered  
; TITLE OF INVENTION: Recombination Sites  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
; STREET: 1100 New York Ave., N. W. Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/005,476

;; FILING DATE: herewith  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/663,002  
;; FILING DATE: 07-JUN-1996  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-2600  
;; TELEFAX: 202-371-2540  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 25 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: both  
;; TOPOLOGY: both  
;; MOLECULE TYPE: cdna  
US-09-005-476-15

Query Match 80.0%; Score 19.2; DB 3; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CAACTTAGTATCAAAAAGCTGAAC 24  
|||||  
Db 24 CAACTTTGTACAAAAAAGCTGAAC 1

RESULT 12  
US-09-233-492-15/c  
; Sequence 15, Application US/09233492  
; Patent No. 6270969  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; TITLE OF INVENTION: Recombinational Cloning Using Engineered  
; TITLE OF INVENTION: Recombination Sites  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C  
; STREET: 1100 New York Ave., N. W. Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/233,492  
; FILING DATE: 20-JAN-1999  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/663,002  
; FILING DATE: 07-JUN-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/486,139  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cdna  
US-09-233-492-15

Query Match 80.0%; Score 19.2; DB 3; Length 25;

Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
Db 24 CAACCTTTGTACAAAAAGCTGAAC 1

## RESULT 13

US-09-296-280-15/c  
; Sequence 15, Application US/09296280  
; Patent No. 6277608  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; APPLICANT: Temple, Gary F.  
; APPLICANT: Fox, Donna K.  
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having  
; FILE REFERENCE: 0942.2850007  
; CURRENT APPLICATION NUMBER: US/09/296,280  
; CURRENT FILING DATE: 1999-04-22  
; EARLIER APPLICATION NUMBER: US 09/177,387  
; EARLIER FILING DATE: 1998-10-23  
; EARLIER APPLICATION NUMBER: US 60/065,930  
; EARLIER FILING DATE: 1997-10-24  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: recombination  
US-09-296-280-15

Query Match 80.0%; Score 19.2; DB 3; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
Db 24 CAACCTTTGTACAAAAAGCTGAAC 1

## RESULT 14

US-09-498-074-15/c  
; Sequence 15, Application US/09498074  
; Patent No. 6534264  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; TITLE OF INVENTION: Recombinational Cloning Using Engineered  
; TITLE OF INVENTION: Recombination Sites  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C  
; STREET: 1100 New York Ave., N. W. Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/498,074  
; FILING DATE: (Herewith)  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/005,476  
; FILING DATE: 12-JAN-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/663,002  
; FILING DATE: 07-JUN-1996  
; CLASSIFICATION:

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/486,139  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cdna  
US-09-498-074-15

Query Match 80.0%; Score 19.2; DB 4; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
Db 24 CAACCTTTGTACAAAAAGCTGAAC 1

## RESULT 15

PCT-US96-10082A-15/c  
; Sequence 15, Application PC/TUS9610082A  
; GENERAL INFORMATION:  
; APPLICANT: Life Technologies, Inc.  
; APPLICANT: 8717 Grovemont Circle  
; APPLICANT: Gaithersburg, MD 20894-9980  
; APPLICANT: United States of America  
; APPLICANT: Brasch, Michael A.  
; TITLE OF INVENTION: Recombinational Cloning Using Engineered  
; TITLE OF INVENTION: Recombination Sites  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C  
; STREET: 1100 New York Ave., N. W. Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10082A  
; FILING DATE: 07-JUN-1996  
; CLASSIFICATION:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cdna  
PCT-US96-10082A-15

Query Match 80.0%; Score 19.2; DB 5; Length 25;

Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 CAACTTAGTATCAAAAAGCTGAAC 24  
| | | | | | | | | | | | | | | | | | | | | |  
Db 24 CAACTTGTACAAAAAAGCTGAAC 1

Search completed: December 3, 2003, 14:51:47  
Job time : 29.562 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 1076.23 seconds  
(without alignments)  
541.991 Million cell updates/sec

Title: US-10-082-772-4  
Perfect score: 24  
Sequence: 1 caacttagtatcaaaaagctgaac 24

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba:\*
  - 2: em\_esthum:\*
  - 3: em\_estin:\*
  - 4: em\_estnu:\*
  - 5: em\_estov:\*
  - 6: em\_estov:\*
  - 7: em\_estro:\*
  - 8: em\_hic:\*
  - 9: gb\_estl:\*
  - 10: gb\_est2:\*
  - 11: gb\_hic:\*
  - 12: gb\_est3:\*
  - 13: gb\_est4:\*
  - 14: gb\_est5:\*
  - 15: em\_estfun:\*
  - 16: em\_estom:\*
  - 17: em\_gss\_hum:\*
  - 18: em\_gss\_inv:\*
  - 19: em\_gss\_pln:\*
  - 20: em\_gss\_vit:\*
  - 21: em\_gss\_fun:\*
  - 22: em\_gss\_mam:\*
  - 23: em\_gss\_mus:\*
  - 24: em\_gss\_pro:\*
  - 25: em\_gss\_rcd:\*
  - 26: em\_gss\_phg:\*
  - 27: em\_gss\_vrl:\*
  - 28: gb\_gss1:\*
  - 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	95.8	664	28	AQ991011
2	22.4	93.3	206	13	BQ156416
3	22.4	93.3	299	13	BY115594
4	22.4	93.3	472	13	BQ157398

5	22.4	93.3	473	13	BQ156404
6	22.4	93.3	597	12	BI422679
7	22.4	93.3	695	28	AQ991039
8	22.4	93.3	712	28	AQ990809
9	22.4	93.3	743	28	AQ990346
10	22.4	93.3	764	28	AQ990110
11	22.4	93.3	769	28	AQ990470
12	21.4	89.2	395	28	AQ991303
13	21.4	89.2	751	28	AQ989566
14	20.8	86.7	672	28	AQ990864
15	20.8	86.7	753	28	AQ990861
16	20.8	86.7	770	28	AQ991774
17	20.8	86.7	791	28	AQ991791
18	20.8	86.7	808	28	AQ990388
19	19.8	82.5	719	28	AQ991352
20	19.8	82.5	756	28	AQ991732
21	19.8	82.5	787	29	BZ971242
22	19.2	80.0	220	9	AI856003
23	19.2	80.0	675	28	AQ991241
24	19.2	80.0	708	28	AQ990869
25	19.2	80.0	746	28	AQ990173
26	18.8	78.3	530	28	AZ263252
27	18.8	78.3	559	28	AZ300017
28	18.8	78.3	653	14	CB422481
29	18.8	78.3	676	29	AG147616
30	18.8	78.3	731	28	AQ975703
31	18.8	78.3	891	29	AG093904
32	18.2	75.8	259	10	BG156704
33	18.2	75.8	388	13	BQ697293
34	18.2	75.8	389	9	AW758911
35	18.2	75.8	391	13	BY510914
36	18.2	75.8	403	9	AW153070
37	18.2	75.8	404	13	BU493456
38	18.2	75.8	407	10	BG317764
39	18.2	75.8	428	9	AI449436
40	18.2	75.8	441	9	AI510342
41	18.2	75.8	449	28	AZ413832
42	18.2	75.8	452	9	AA120033
43	18.2	75.8	472	12	BM188477
44	18.2	75.8	489	29	AG256781
45	18.2	75.8	500	28	BH885061

ALIGNMENTS

RESULT 1  
AQ991011  
LOCUS  
DEFINITION

664 bp DNA linear GSS 14-AUG-2000  
Rfc01864 Photorhabdus luminescens strain W14 M13 library  
Phototaxidus luminescens genomic clone PLG01864, genomic survey  
sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AQ991011  
AQ991011.1  
G1:9649605  
GSS.  
Phototaxidus luminescens

Phototaxidus luminescens  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Phototaxidus.

REFERENCE  
AUTHORS

1 (bases 1 to 664)  
ffrench-Constant R.H., Waterfield N., Burland V., Perna N.T.,  
Daborn P.J., Bowen D. and Blattner F.R.

A genomic sample sequence of the entomopathogenic bacterium  
Phototaxidus luminescens W14: potential implications for virulence

JOURNAL  
MEDLINE  
PUBMED

20378633  
10919786

COMMENT

Contact: ffrench-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK  
Tel: (44) 1225 826621  
Fax: (44) 1225 826779

Email: bsrfc@bath.ac.uk  
 This is one of 2,122 random reads from the M13 library. For  
 annotation of identified clones (BLASTX, BLASTN and mapping to E.  
 coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic  
 Acids Res.  
 Seq primer: M13 Forward  
 Class: shotgun.

#### FEATURES

source  
 Location/Qualifiers  
 1..664  
 /organism="Photorhabdus luminescens"  
 /mol\_type="genomic DNA"  
 /strain="W14"  
 /db\_xref="taxon:29488"  
 /clone="PLG01864"  
 /dev\_stage="primary phase variant"  
 /clone\_lib="Photorhabdus luminescens strain W14 M13  
 library"  
 /notes="genomic DNA from strain W14 was size selected (1-2  
 kb) and then cloned into M13 Janus."  
 BASE COUNT 178 a 146 c 162 g 175 t 3 others  
 ORIGIN

Query Match 95.8%; Score 23; DB 28; Length 664;  
 Best Local Similarity 95.8%; Pred. No. 30;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAACCTAGTATCAAAAAGCTGAAC 24  
 |||||  
 Db 615 CAACCTAGTATCAAAAAGCTGAAC 638

#### RESULT 2

LOCUS B0156416 206 bp mRNA linear EST 24-APR-2002  
 DEFINITION NF092F02IR1F1027 Irradiated Medicago truncatula cDNA clone  
 B0156416

ACCESSION B0156416  
 VERSION B0156416.1 GI:20293475  
 KEYWORDS EST  
 SOURCE Medicago truncatula (barrel medic)  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

REFERENCE 1 (bases 1 to 206)  
 Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,  
 Flores, H.R., Iman, J.T., Weller, J.W. and May, G.D.  
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
 Medicago truncatula irradiated library  
 Unpublished

JOURNAL Contact: May GD  
 COMMENT Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7391  
 Fax: 580 221 7380

Email: gdmay@noble.org  
 Insert Length: 206 Std Error: 0.00  
 Plate: 092 row: F column: 02  
 Seq primer: TCACACGGAACAGCTGAC.

#### FEATURES

source  
 Location/Qualifiers  
 1..206  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3880"  
 /clone="NF092F02IR"  
 /tissue\_type="seedlings"  
 /dev\_stage="seedling"  
 /clone\_lib="Irradiated"  
 /notes="Vector: Lambda zap; Seedlings were exposed either  
 to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.  
 Gamma-irradiated samples were harvested at 6, 12, 24 and

48 hours after treatment. UV-irradiated samples were  
 harvested 24 hours post-treatment. cDNA was prepared from  
 polyA+ enriched, pooled samples of equivalent amounts of  
 total RNA from each sample. The cDNA was directionally  
 ligated into the Uni-Zap XR vector (Stratagene) and  
 packaged using the Gigapack III Gold packaging extracts.  
 Phagemids containing cDNA inserts were in vivo excised  
 from the recombinant Uni-Zap XR vector using ExAestit  
 helper phage and the E. coli strain XL1-Blue MRF'  
 (Stratagene). Excised plasmids were plated using SOLR  
 cells."

BASE COUNT 81 a 27 c 39 g 59 t  
 ORIGIN

Query Match 93.3%; Score 22.4; DB 13; Length 206;  
 Best Local Similarity 95.8%; Pred. No. 37;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAACCTAGTATCAAAAAGCTGAAC 24  
 |||||  
 Db 144 CAACCTAGTATCAAAAAGCTGAAC 167

#### RESULT 3

LOCUS B0115594/c

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

B0115594 299 bp mRNA linear EST 08-DEC-2002  
 B0115594 RIKEN full-length enriched, 18 days embryo whole body Mus  
 musculus cDNA clone L430040C03 5', mRNA sequence.

B0115594

B0115594.1 GI:26226695

EST.

Mus musculus (house mouse)

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 299)

Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.

, Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,

Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,

Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,

Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani,

, L.E., Cousins, S., Dalla, S., Dragani, T.A., Fletcher, C.F., Forrest,

, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.

, Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,

Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,

King, B.L., Kongay, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons

, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki

, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Petosa, G.,

Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,

Ramachandran, S., Ravasi, T., Reed, J.C., Read, D.J., Reid, J., Ring

, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou

, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale

, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,

Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa

, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,

Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura

, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.

, Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii

, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata

, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander

, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane

, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komno, H., Miyazaki, A.

, Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.

, Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami

, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct

Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multichannel sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

#### FEATURES

source  
Location/Qualifiers

1. .239

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CS7BL/6J"

/db\_xref="taxon:10090"

/clone="I430040C03"

/tissue\_type="whole body"

/dev\_stage="18 days embryo"

/clone\_lib="RIKEN full-length enriched, 18 days embryo

whole body"

85 a 50 c 54 g 110 t

BASE COUNT  
ORIGIN

Query Match

Best Local Similarity 93.3%; Score 22.4; DB 13; Length 299;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAACTTAGTATCAAAAAGCTGAAC 24

|||||

Db 269 CAACTTAGTATCAAAAAGCTGAAC 246

RESULT 4

LOCUS

DEFINITION

NF104D071R1F1062 Irradiated Medicago truncatula cDNA clone

NF104D071R 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Medicago truncatula (barrel medic)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

REFERENCE

1 (bases 1 to 472)

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,

Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula irradiated library

Unpublished

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7380

Email: [gdmay@noble.org](mailto:gdmay@noble.org)

Insert Length: 472 Std Error: 0.00

Plate: 104 row: D column: 07

Seq primer: TCACACAGGAACAGCTATGAC.

Location/Qualifiers

1. .472

/organism="Medicago truncatula"

/mol\_type="mRNA"

/db\_xref="taxon:3880"

/clone="NF104D071R"

/tissue\_type="seedlings"

/dev\_stage="seedling"

/clone\_lib="Irradiated"

/notes="Vector: Lambda Zap; Seedlings were exposed either

to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.

Gamma-irradiated samples were harvested at 6, 12, 24 and

48 hours after treatment. UV-irradiated samples were

harvested 24 hours post-treatment. cDNA was prepared from

polyA+ enriched, pooled samples of equivalent amounts of

total RNA from each sample. The cDNA was directionally

ligated into the Uni-Zap XR vector (Stratagene) and

packaged using the Gigapack III Gold packaging extracts.

Phagemids containing cDNA inserts were in vivo excised

from the recombinant Uni-Zap XR vector using ExAssist

helper phage and the E. coli strain XL1-Blue MRF.

(Stratagene). Excised plasmids were plated using SOLR

cells."

BASE COUNT 162 a 89 c 95 g 126 t

ORIGIN

Query Match

Best Local Similarity 93.3%; Score 22.4; DB 13; Length 472;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAACTTAGTATCAAAAAGCTGAAC 24

|||||

Db 357 CAACTTAGTATCAAAAAGCTGAAC 380

RESULT 5

LOCUS

DEFINITION

NF092E031R1F1023 Irradiated Medicago truncatula cDNA clone

NF092E031R 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Medicago truncatula (barrel medic)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

REFERENCE

1 (bases 1 to 473)

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,

Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula irradiated library

Unpublished

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7380

Email: [gdmay@noble.org](mailto:gdmay@noble.org)

Insert Length: 473 Std Error: 0.00

Plate: 092 row: E column: 03

Seq primer: TCACACAGGAACAGCTATGAC.

Location/Qualifiers

source

```

1..473
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF092031R"
/tissue_type="seedlings"
/dev_stage="seedling"
/clone_lib="Irradiated"
/notes="Vector: Lambda Zap; Seedlings were exposed either
to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
Gamma-irradiated samples were harvested at 6, 12, 24 and
48 hours after treatment. UV-irradiated samples were
harvested 24 hours post-treatment. cDNA was prepared from
polyA+ enriched, pooled samples of equivalent amounts of
total RNA from each sample. The cDNA was directionally
ligated into the Uni-Zap XR vector (Stratagene) and
packaged using the Gigapack III Gold packaging extracts.
Phagemids containing cDNA inserts were in vivo excised
from the recombinant Uni-Zap XR vector using ExAssist
helper phage and the E. coli strain XL1-Blue MRF'
(Stratagene). Excised plasmids were plated using SOLR
cells."
BASE COUNT      162 a  90 c  95 g  126 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 13; Length 473;
Best Local Similarity 95.8%; Pred. No. 48;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CAACTTAGTATCAAAAAGCTGAAC 24
|||||
Db 358 CAACTTAGTATCAAAAAGCTGAAC 381
|||||

```

```

RESULT 6
BI422679
LOCUS
DEFINITION
EST533345 tomato callus, TAMU Lycopersicon esculentum cDNA clone
c1EC71G2 5' end, mRNA sequence.
BI422679
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 597)
Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E.,
Liang, F., Upton, J., Craven, M.B., Bowman, C.L., Ahn, S., Renning,
C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
Generation of ESTs from tomato callus tissue
Unpublished
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1..597
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="c1EC71G2"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XL1-Blue MRF'"
/clone_lib="tomato callus, TAMU"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni laboratory; c1EC - Coryedons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.

```

source

```

BASE COUNT      193 a  109 c  131 g  164 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 12; Length 597;
Best Local Similarity 95.8%; Pred. No. 51;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CAACTTAGTATCAAAAAGCTGAAC 24
|||||
Db 616 CAACTTAGTATCAAAAAGCTGAAC 639
|||||
RESULT 8

```

```

BASE COUNT      193 a  109 c  131 g  164 t
ORIGIN
Query Match      93.3%; Score 22.4; DB 12; Length 597;
Best Local Similarity 95.8%; Pred. No. 51;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CAACTTAGTATCAAAAAGCTGAAC 24
|||||
Db 529 CAACTTAGTATCAAAAAGCTGAAC 552
|||||
RESULT 7
LOCUS
DEFINITION
Photorhabdus luminescens genomic clone PLG01894, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Photorhabdus luminescens
Photorhabdus luminescens
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Photorhabdus.
1 (bases 1 to 695)
ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,
Daborn, P.J., Bowen, D. and Blattner, P.R.
A genomic sample sequence of the entomopathogenic bacterium
Photorhabdus luminescens W14: potential implications for virulence
Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
20378633
PUBMED
10919786
COMMENT
Contact: ffrench-Constant RH
Department of Biology and Biochemistry
University of Bath
South Building, Bath BA2 7AY, UK
Tel: (44) 1225 826621
Fax: (44) 1225 826779
Email: bssrfc@bath.ac.uk
This is one of 2,122 random reads from the M13 library. For
annotation of identified clones (BLASTX, BLASTN and mapping to E.
coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic
Acids Res.
Seq primer: M13 Forward
Class: shotgun.
Location/Qualifiers
1..695
/organism="Photorhabdus luminescens"
/mol_type="genomic DNA"
/strain="W14"
/db_xref="taxon:29488"
/dev_stage="primary phase variant"
/clone="PLG01894"
/clone_lib="Photorhabdus luminescens strain W14 M13
library"
/notes="Genomic DNA from strain W14 was size selected (1-2
kb) and then cloned into M13 Janus."

```

```

BASE COUNT      193 a  148 c  165 g  187 t  2 others
ORIGIN
Query Match      93.3%; Score 22.4; DB 28; Length 695;
Best Local Similarity 95.8%; Pred. No. 54;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CAACTTAGTATCAAAAAGCTGAAC 24
|||||
Db 616 CAACTTAGTATCAAAAAGCTGAAC 639
|||||
RESULT 8

```

<p><b>AQ990809</b> LOCUS DEFINITION</p>		<p>712 bp DNA linear GSS 14-AUG-2000 Rf01638 Photorhabdus luminescens strain W14 M13 library Photorhabdus luminescens genomic clone PLG01638, genomic survey sequence.</p>	
<p><b>ACCESSION</b> VERSION KEYWORDS SOURCE ORGANISM</p>		<p>AQ990809 AQ990809.1 GI:9649403 GSS. Photorhabdus luminescens Photorhabdus luminescens Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus. 1 (bases 1 to 712)</p>	
<p><b>REFERENCE</b> AUTHORS</p>		<p>french-Constant R.H., Waterfield,N., Burland,V., Perna,N.T., Daborn,P.J., Bowen,D. and Blattner,F.R. A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)</p>	
<p><b>TITLE</b></p>		<p>A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)</p>	
<p><b>JOURNAL</b> MEDLINE PUBMED COMMENT</p>		<p>20378633 10919786 Contact: french-Constant RH Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779 Email: bsrfc@bath.ac.uk This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res. Seq primer: M13 Forward Class: shotgun.</p>	
<p><b>FEATURES</b> source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 218 a 144 c 163 g 187 t 2 others</p>	
<p><b>BASE COUNT</b> ORIGIN</p>		<p>218 a 144 c 163 g 187 t 2 others</p>	
<p><b>FEATURES</b> source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 218 a 144 c 163 g 187 t 2 others</p>	
<p><b>BASE COUNT</b> ORIGIN</p>		<p>218 a 144 c 163 g 187 t 2 others</p>	
<p><b>FEATURES</b> source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 218 a 144 c 163 g 187 t 2 others</p>	
<p><b>BASE COUNT</b> ORIGIN</p>		<p>218 a 144 c 163 g 187 t 2 others</p>	
<p><b>FEATURES</b> source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 218 a 144 c 163 g 187 t 2 others</p>	
<p><b>BASE COUNT</b> ORIGIN</p>		<p>218 a 144 c 163 g 187 t 2 others</p>	
<p><b>FEATURES</b> source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 218 a 144 c 163 g 187 t 2 others</p>	
<p><b>BASE COUNT</b> ORIGIN</p>		<p>218 a 144 c 163 g 187 t 2 others</p>	
<p><b>FEATURES</b> source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 218 a 144 c 163 g 187 t 2 others</p>	
<p><b>BASE COUNT</b> ORIGIN</p>		<p>218 a 144 c 163 g 187 t 2 others</p>	
<p><b>FEATURES</b> source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 218 a 144 c 163 g 187 t 2 others</p>	
<p><b>BASE COUNT</b> ORIGIN</p>		<p>218 a 144 c 163 g 187 t 2 others</p>	
<p><b>FEATURES</b> source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 218 a 144 c 163 g 187 t 2 others</p>	
<p><b>BASE COUNT</b> ORIGIN</p>		<p>218 a 144 c 163 g 187 t 2 others</p>	
<p><b>FEATURES</b> source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 218 a 144 c 163 g 187 t 2 others</p>	
<p><b>BASE COUNT</b> ORIGIN</p>		<p>218 a 144 c 163 g 187 t 2 others</p>	
<p><b>FEATURES</b> source</p>		<p>Location/Qual</p>	

<p><b>AQ990809</b> LOCUS DEFINITION</p>		<p>712 bp DNA linear GSS 14-AUG-2000 Rf01638 Photorhabdus luminescens strain W14 M13 library Photorhabdus luminescens genomic clone PLG01638, genomic survey sequence.</p>	
<p>ACCESSION VERSION KEYWORDS SOURCE ORGANISM</p>		<p>AQ990809 AQ990809.1 GI:9649403 GSS. Photorhabdus luminescens Photorhabdus luminescens Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus. 1 (bases 1 to 712)</p>	
<p>REFERENCE AUTHORS</p>		<p>french-Constant R.H., Waterfield,N., Burland,V., Perna,N.T., Daborn,P.J., Bowen,D. and Blattner,F.R. A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)</p>	
<p>TITLE</p>		<p>A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)</p>	
<p>JOURNAL MEDLINE PUBMED COMMENT</p>		<p>20378633 10919786 Contact: french-Constant RH Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779 Email: bsrfc@bath.ac.uk This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res. Seq primer: M13 Forward Class: shotgun.</p>	
<p>FEATURES source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 214 a 144 c 163 g 187 t 2 others</p>	
<p>BASE COUNT ORIGIN</p>		<p>218 a 144 c 163 g 187 t</p>	
<p>Query Match Best Local Similarity Matches</p>		<p>93.3%; Score 22.4; DB 28; Length 712; 95.8%; Pred. No. 54; 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</p>	
<p>DB</p>		<p>1 CAACCTAGTATCAAAAAGCTGAAC 24       430 CAACCTAGTATCAAAAAGCTGAAC 453</p>	
<p>RESULT 9 LOCUS DEFINITION</p>		<p>AQ990346 Rf01106 Photorhabdus luminescens strain W14 M13 library Photorhabdus luminescens genomic clone PLG01106, genomic survey sequence.</p>	
<p>ACCESSION VERSION KEYWORDS SOURCE ORGANISM</p>		<p>AQ990346 AQ990346.1 GI:9648940 GSS. Photorhabdus luminescens Photorhabdus luminescens Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus. 1 (bases 1 to 743)</p>	
<p>REFERENCE AUTHORS</p>		<p>french-Constant R.H., Waterfield,N., Burland,V., Perna,N.T., Daborn,P.J., Bowen,D. and Blattner,F.R. A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)</p>	
<p>TITLE</p>		<p>A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)</p>	
<p>JOURNAL MEDLINE PUBMED COMMENT</p>		<p>20378633 10919786 Contact: french-Constant RH Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779 Email: bsrfc@bath.ac.uk This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res. Seq primer: M13 Forward Class: shotgun.</p>	
<p>FEATURES source</p>		<p>Location/Qualifiers 1..712 /organism="Photorhabdus luminescens" /mol_type="Genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01638" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 214 a 158 c 169 g 200 t 2 others</p>	
<p>BASE COUNT ORIGIN</p>		<p>214 a 158 c 169 g 200 t</p>	
<p>Query Match Best Local Similarity Matches</p>		<p>93.3%; Score 22.4; DB 28; Length 743; 95.8%; Pred. No. 55; 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</p>	
<p>DB</p>		<p>1 CAACCTAGTATCAAAAAGCTGAAC 24       603 CAACCTAGTATCAAAAAGCTGAAC 626</p>	
<p>RESULT 10 LOCUS DEFINITION</p>		<p>AQ990110 Rf00827 Photorhabdus luminescens strain W14 M13 library Photorhabdus luminescens genomic clone PLG00827, genomic survey sequence.</p>	
<p>ACCESSION VERSION KEYWORDS SOURCE ORGANISM</p>		<p>AQ990110 AQ990110.1 GI:9648704 GSS. Photorhabdus luminescens Photorhabdus luminescens Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus. 1 (bases 1 to 764)</p>	
<p>REFERENCE AUTHORS</p>		<p>french-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T., Daborn,P.J., Bowen,D. and Blattner,F.R. A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)</p>	
<p>TITLE</p>		<p>A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)</p>	
<p>JOURNAL MEDLINE PUBMED COMMENT</p>		<p>20378633 10919786 Contact: french-Constant RH Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779 Email: bsrfc@bath.ac.uk This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res.</p>	

Acids Res. Seq primer: M13 Forward Class: shotgun.		Acids Res. Seq primer: M13 Forward Class: shotgun.	
FEATURES source	Location/Qualifiers	FEATURES source	Location/Qualifiers
	1..764		1..395
	/organism="Photorhabdus luminescens"		/organism="Photorhabdus luminescens"
	/mol_type="genomic DNA"		/mol_type="genomic DNA"
	/strain="W14"		/strain="W14"
BASE COUNT ORIGIN	/db_xref="taxon:29488"	BASE COUNT ORIGIN	/db_xref="taxon:29488"
	/clones="PLG00827"		/clones="PLG02205"
	/dev_stage="primary phase variant"		/dev_stage="primary phase variant"
	/clone_lib="Photorhabdus luminescens strain W14 M13 library"		/clone_lib="Photorhabdus luminescens strain W14 M13 library"
	/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."		/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."
215 a 170 c 171 g 203 t 5 others		135 a 72 c 63 g 101 t 24 others	
Query Match Best Local Similarity 93.3%; Score 22.4; DB 28; Length 764; Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		Query Match Best Local Similarity 89.2%; Score 21.4; DB 28; Length 395; Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1 CAACCTAGTATCAAAAAGCTGAAC 24	QY	1 CAACCTAGTATCAAAAAGCTGAAC 24
Db	698 CAACCTAGTATCAAAAAGCTGAAC 721	Db	187 CAACCTAGTATCAAAAAGCTGAAC 210
RESULT 11		RESULT 12	
AQ990470		AQ991303	
LOCUS		LOCUS	
DEFINITION		DEFINITION	
Photorhabdus luminescens strain W14 M13 library		Photorhabdus luminescens strain W14 M13 library	
sequence.		sequence.	
ACCESSION		ACCESSION	
AQ990470.1 GI:9649064		AQ991303.1 GI:9649897	
VERSION		VERSION	
GSS.		GSS.	
KEYWORDS		KEYWORDS	
SOURCE		SOURCE	
ORGANISM		ORGANISM	
Photorhabdus luminescens		Photorhabdus luminescens	
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	
Enterobacteriaceae; Photorhabdus.		Enterobacteriaceae; Photorhabdus.	
1 (bases 1 to 769)		1 (bases 1 to 395)	
french-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,		french-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,	
Daborn,P.J., Bowen,D. and Blattner,F.R.		Daborn,P.J., Bowen,D. and Blattner,F.R.	
A genomic sample sequence of the entomopathogenic bacterium		A genomic sample sequence of the entomopathogenic bacterium	
Photorhabdus luminescens W14; potential implications for virulence		Photorhabdus luminescens W14; potential implications for virulence	
Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)		Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)	
20378633		20378633	
PUBMED		PUBMED	
10919786		10919786	
COMMENT		COMMENT	
Contact: french-Constant RH		Contact: french-Constant RH	
Department of Biology and Biochemistry		Department of Biology and Biochemistry	
University of Bath		University of Bath	
South Building, Bath BA2 7AY, UK		South Building, Bath BA2 7AY, UK	
Tel: (44) 1225 826621		Tel: (44) 1225 826621	
Fax: (44) 1225 826779		Fax: (44) 1225 826779	
Email: bssrfc@bath.ac.uk		Email: bssrfc@bath.ac.uk	
This is one of 2,122 random reads from the M13 library. For		This is one of 2,122 random reads from the M13 library. For	
annotation of identified clones (BLASTX, BLASTN and mapping to E.		annotation of identified clones (BLASTX, BLASTN and mapping to E.	
coli K12 genome) please see french-Constant et al. 2000, Nucleic		coli K12 genome) please see french-Constant et al. 2000, Nucleic	
Acids Res.		Acids Res.	
Seq primer: M13 Forward		Seq primer: M13 Forward	
Class: shotgun.		Class: shotgun.	
Location/Qualifiers		Location/Qualifiers	
1..769		1..395	
/organism="Photorhabdus luminescens"		/organism="Photorhabdus luminescens"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
/strain="W14"		/strain="W14"	
/db_xref="taxon:29488"		/db_xref="taxon:29488"	
/clones="PLG01245"		/clones="PLG02205"	
/dev_stage="primary phase variant"		/dev_stage="primary phase variant"	
/clone_lib="Photorhabdus luminescens strain W14 M13 library"		/clone_lib="Photorhabdus luminescens strain W14 M13 library"	
/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."		/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."	

JOURNAL	Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
MEDLINE	20378633
PUBMED	10919786
COMMENT	Contact: french-Constant RH Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779 Email: bsarfc@bath.ac.uk This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res. Seq primer: M13 Forward Class: shotgun.
FEATURES	Location/Qualifiers 1..672 /organism="Photorhabdus luminescens" /mol_type="genomic DNA" /strain="W14" /db_xref="taxon:29488" /clone="PLG01701" /dev_stage="primary phase variant" /clone_lib="Photorhabdus luminescens strain W14 M13 library" /note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus." 184 a 146 c 161 g 179 t 2 others
BASE COUNT	
ORIGIN	
Query Match	86.7%; Score 20.8; DB 28; Length 672;
Best Local Similarity	91.7%; Pred. No. 2.5e+02;
Matches	22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	1 CAACTTAGTATCAAAAAGCTGAAC 24
Db	614 CCACTTAGTATCAAAAAGCTGAAC 637
RESULT 15	
AQ990861	
LOCUS	AQ990861 753 bp DNA linear GSS 14-AUG-2000
DEFINITION	Rfc01698 Photorhabdus luminescens strain W14 M13 library Photorhabdus luminescens genomic clone PLG01698, genomic survey sequence.
ACCESSION	AQ990861
VERSION	AQ990861.1 GI:9649455
KEYWORDS	GSS.
SOURCE	Photorhabdus luminescens
ORGANISM	Photorhabdus luminescens Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus.
REFERENCE	1 (bases 1 to 753) french-Constant R.H., Waterfield,N., Burland,V., Perna,N.T., Daborn,P.J., Bowen,D. and Blattner,F.R. A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence
AUTHORS	
TITLE	
JOURNAL	Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
MEDLINE	20378633
PUBMED	10919786
COMMENT	Contact: french-Constant RH Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779 Email: bsarfc@bath.ac.uk This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res. Seq primer: M13 Forward

FEATURES Class: shotgun, Location/Qualifiers  
source 1..753  
/organism="Photorhabdus luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG01698"  
/dev\_stage="primary phase variant"  
/clone\_lib="Photorhabdus luminescens strain W14 M13 library"  
/notes="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."  
BASE COUNT 216 a 160 c 169 g 203 t 5 others  
ORIGIN  
Query Match 86.7%; Score 20.8; DB 28; Length 753;  
Best Local Similarity 91.7%; Pred. NO. 2.6e+02;  
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
Db 615 CCACCTTAGTATCAAAAAGCTGAAC 638

Search completed: December 3, 2003, 14:49:19  
JOB time : 1076.23 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:45 ; Search time 28.562 Seconds  
(without alignments)  
370.884 Million cell updates/sec

Title: US-10-082-772-3

Perfect score: 24

Sequence: 1 gttcagctttttgatactaagtgtg 24

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22.4	93.3	201	1	US-08-021-667A-18
2	22.4	93.3	201	1	US-08-410-584A-18
3	22.4	93.3	201	1	US-08-728-785A-18
4	22.4	93.3	4909	3	US-08-556-978B-78
5	22.4	93.3	7652	1	US-07-590-988A-1
6	21.2	88.3	25	3	US-09-295-280-43
7	20.4	85.0	25	3	US-09-295-280-42
8	19.4	80.8	21	3	US-08-894-511-7
9	19.4	80.8	21	4	US-09-655-728-7
10	19.2	80.0	25	3	US-09-233-493-15
11	19.2	80.0	25	3	US-09-005-476-15
12	19.2	80.0	25	3	US-09-233-492-15
13	19.2	80.0	25	3	US-09-295-280-15
14	19.2	80.0	25	4	US-09-498-074-15
15	19.2	80.0	25	5	PCT-US96-10082A-15
16	18.6	77.5	25	3	US-09-233-493-5
17	18.6	77.5	25	3	US-09-005-476-5
18	18.6	77.5	25	3	US-09-233-492-5
19	18.6	77.5	25	3	US-09-295-280-5
20	18.6	77.5	25	4	US-09-498-074-5
21	18.6	77.5	25	5	PCT-US96-10082A-5
22	18.4	76.7	25	3	US-09-295-280-39
23	18.2	75.8	580073	4	US-08-545-528D-1
24	17.8	74.2	25	3	US-09-295-280-41
25	17.8	74.2	228	4	US-09-107-532A-667
26	17.8	74.2	1123	3	US-09-188-930-28
27	17.8	74.2	1123	3	US-09-188-930-203

28	17.8	74.2	1123	4	US-09-312-283C-28	Sequence 28, Appl
29	17.8	74.2	1123	4	US-09-312-283C-203	Sequence 203, Appl
30	17.8	74.2	2408	1	US-08-486-013-69	Sequence 69, Appl
31	17.8	74.2	2408	2	US-08-482-279-69	Sequence 69, Appl
32	17.8	74.2	2408	2	US-08-342-268-69	Sequence 69, Appl
33	17.8	74.2	2408	3	US-09-015-968-69	Sequence 69, Appl
34	17.8	74.2	2408	4	US-09-397-386-69	Sequence 69, Appl
35	17.8	74.2	3484	4	US-09-308-090-1	Sequence 1, Appl
36	17.8	74.2	3484	4	US-09-380-090A-1	Sequence 1, Appl
37	17.8	74.2	3757	2	US-09-016-366A-13	Sequence 13, Appl
38	17.8	74.2	3757	2	US-08-978-404B-19	Sequence 19, Appl
39	17.8	74.2	5349	4	US-09-068-101-7	Sequence 7, Appl
40	17.8	74.2	5611	4	US-09-068-101-10	Sequence 10, Appl
41	17.6	73.3	25	3	US-09-233-493-9	Sequence 9, Appl
42	17.6	73.3	25	3	US-09-233-493-11	Sequence 11, Appl
43	17.6	73.3	25	3	US-09-233-493-16	Sequence 16, Appl
44	17.6	73.3	25	3	US-09-005-476-9	Sequence 9, Appl
45	17.6	73.3	25	3	US-09-005-476-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1  
US-08-021-667A-18  
; Sequence 18, Application US/08021667A  
; Patent No. 5434049  
; GENERAL INFORMATION:  
; APPLICANT: Okano, Kazunori  
; APPLICANT: Kambata, Hideki  
; TITLE OF INVENTION: POLYNUCLEOTIDE CAPTURING TIP AND  
; TITLE OF INVENTION: POLYNUCLEOTIDE PREPARATIVE METHOD AND DETECTION  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Antonelli, Terry, Stout & Kraus  
; STREET: Suite 600, 1919 Pennsylvania Ave., NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/021,667A  
; FILING DATE: 19930224  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Terry, David T.  
; REGISTRATION NUMBER: 20,178  
; REFERENCE/DOCKET NUMBER: 520.3:930X00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-828-0300  
; TELEFAX: 202-828-0380  
; TELEX: 440280/248545  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 201 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: YES  
; ANTI-SENSE: NO  
US-08-021-667A-18

Query Match 93.3%; Score 22.4; DB 1; Length 201;  
Best Local Similarity 95.8%; Pred. No. 0.28;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
 |||||  
 Db 40 GTTCAGCTTTTGTACTAAGTTG 63

## RESULT 2

US-08-410-544-18  
 ; Sequence 18, Application US/08410544  
 ; Patent No. 5607646  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Okano, Kazunori  
 ; APPLICANT: Kambara, Hideki  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE CAPTURING TIP AND  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE PREPARATIVE METHOD AND DETECTION  
 ; TITLE OF INVENTION: METHOD USING SAME  
 ; NUMBER OF SEQUENCES: 18  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Antonelli, Terry, Stout & Kraus  
 ; STREET: Suite 600, 1919 Pennsylvania Ave., NW  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION NUMBER: US/08/410,544  
 ; FILING DATE: 24-FEB-1993  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/021,667  
 ; FILING DATE: 24-FEB-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Terry, David T.  
 ; REGISTRATION NUMBER: 20,178  
 ; REFERENCE/DOCKET NUMBER: 520.31930X00  
 ; TELEPHONE: 202-828-0300  
 ; TELEFAX: 202-828-0380  
 ; TELEX: 248545  
 ; INFORMATION FOR SEQ ID NO: 18:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 201 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHETICAL: YES  
 ; ANTI-SENSE: NO  
 ; US-08-410-544-18

Query Match 93.3%; Score 22.4; DB 1; Length 201;  
 Best Local Similarity 95.8%; Pred. No. 0.28;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
 |||||  
 Db 40 GTTCAGCTTTTGTACTAAGTTG 63

## RESULT 3

US-08-728-785A-18  
 ; Sequence 18, Application US/08728785A  
 ; Patent No. 5817506  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Okano, Kazunori  
 ; APPLICANT: Kambara, Hideki  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE CAPTURING TIP AND  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE PREPARATIVE METHOD AND DETECTION  
 ; TITLE OF INVENTION: METHOD USING SAME

Query Match 93.3%; Score 22.4; DB 1; Length 201;  
 Best Local Similarity 95.8%; Pred. No. 0.28;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
 |||||  
 Db 40 GTTCAGCTTTTGTACTAAGTTG 63

## RESULT 4

US-08-556-978B-78/c  
 ; Sequence 78, Application US/08556978B  
 ; Patent No. 6268169  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FARNESTOCK, STEPHEN P.  
 ; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED  
 ; TITLE OF INVENTION: SPIDER SILK ANALOGS  
 ; NUMBER OF SEQUENCES: 107  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
 ; STREET: 1007 MARKET STREET  
 ; CITY: WILMINGTON  
 ; STATE: DELAWARE  
 ; COUNTRY: UNITED STATES OF AMERICA  
 ; ZIP: 19898  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: DISKETTE, 3.50 INCH  
 ; COMPUTER: IBM PC COMPATIBLE  
 ; OPERATING SYSTEM: MICROSOFT WINDOWS 95  
 ; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/556,978B

Query Match 93.3%; Score 22.4; DB 1; Length 201;  
 Best Local Similarity 95.8%; Pred. No. 0.28;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
 |||||  
 Db 40 GTTCAGCTTTTGTACTAAGTTG 63

## RESULT 4

US-08-556-978B-78/c  
 ; Sequence 78, Application US/08556978B  
 ; Patent No. 6268169  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FARNESTOCK, STEPHEN P.  
 ; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED  
 ; TITLE OF INVENTION: SPIDER SILK ANALOGS  
 ; NUMBER OF SEQUENCES: 107  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
 ; STREET: 1007 MARKET STREET  
 ; CITY: WILMINGTON  
 ; STATE: DELAWARE  
 ; COUNTRY: UNITED STATES OF AMERICA  
 ; ZIP: 19898  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: DISKETTE, 3.50 INCH  
 ; COMPUTER: IBM PC COMPATIBLE  
 ; OPERATING SYSTEM: MICROSOFT WINDOWS 95  
 ; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/556,978B

;; FILING DATE: 435  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/077,600  
;; FILING DATE: JUNE 15, 1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: FLOYD, LINDA AXAMETHY  
;; REGISTRATION NUMBER: 33,692  
;; REFERENCE/DOCKET NUMBER: CR-9389-A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 302-892-8112  
;; TELEFAX: 302-773-0164  
;; INFORMATION FOR SEQ ID NO: 78:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4909 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
US-08-556-978B-78  
  
Query Match 93.3%; Score 22.4; DB 3; Length 4909;  
Best Local Similarity 95.8%; Pred. No. 0.32; 1; Indels 0; Gaps 0;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||  
DB 2378 GTTCAGCTTTTGTACTAAGTTG 2355  
|||||  
  
RESULT 5  
US-07-590-988A-1  
; Sequence 1, Application US/07590988A  
; Patent No. 5227288  
; GENERAL INFORMATION:  
; APPLICANT: Blattner, Frederick R.  
; TITLE OF INVENTION: DNA Sequencing Vector with  
; TITLE OF INVENTION: Reversible Insert  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles and Brady  
; STREET: P.O. BOX 2113  
; STREET: FIRST WISCONSIN PLAZA  
; CITY: MADISON  
; STATE: WISCONSIN  
; COUNTRY: U.S.A.  
; ZIP: 53701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette - 3.50 inch, 800Kb storage  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh  
; SOFTWARE: Microsoft Word 4.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/590,988A  
; FILING DATE: 19901001  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Nicholas J. Seay  
; REGISTRATION NUMBER: 27,386  
; REFERENCE/DOCKET NUMBER: 9629691682  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (608) 251-5000  
; TELEFAX: (608) 251-9166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7652 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: Other nucleic acid;  
; DESCRIPTION: synthetic recombinant plasmid  
; HYPOTHETICAL: no  
; ANTI-SENSE: no

US-07-590-988A-1  
  
Query Match 93.3%; Score 22.4; DB 1; Length 7652;  
Best Local Similarity 95.8%; Pred. No. 0.33; 1; Indels 0; Gaps 0;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||  
DB 6071 GTTCAGCTTTTGTACTAAGTTG 6094  
|||||  
  
RESULT 6  
US-09-296-280-43  
; Sequence 43, Application US/09296280  
; Patent No. 6277608  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; APPLICANT: Temple, Gary F.  
; APPLICANT: Fox, Donna K.  
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having  
; TITLE OF INVENTION: Recombination Sites  
; FILE REFERENCE: 0942.2850007  
; CURRENT APPLICATION NUMBER: US/09/296,280  
; CURRENT FILING DATE: 1999-04-22  
; EARLIER APPLICATION NUMBER: US 09/177,387  
; EARLIER FILING DATE: 1998-10-23  
; EARLIER APPLICATION NUMBER: US 60/065,930  
; EARLIER FILING DATE: 1997-10-24  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 43  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: recombination  
; OTHER INFORMATION: Products  
US-09-296-280-43  
  
Query Match 88.3%; Score 21.2; DB 3; Length 25;  
Best Local Similarity 83.3%; Pred. No. 0.87; 1; Indels 0; Gaps 0;  
Matches 20; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||  
DB 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||  
  
RESULT 7  
US-09-296-280-42  
; Sequence 42, Application US/09296280  
; Patent No. 6277608  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; APPLICANT: Temple, Gary F.  
; APPLICANT: Fox, Donna K.  
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having  
; TITLE OF INVENTION: Recombination Sites  
; FILE REFERENCE: 0942.2850007  
; CURRENT APPLICATION NUMBER: US/09/296,280  
; CURRENT FILING DATE: 1999-04-22  
; EARLIER APPLICATION NUMBER: US 09/177,387  
; EARLIER FILING DATE: 1998-10-23  
; EARLIER APPLICATION NUMBER: US 60/065,930  
; EARLIER FILING DATE: 1997-10-24  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 42  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Unknown

FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: recombination  
OTHER INFORMATION: products

US-09-296-280-42

Query Match 85.0%; Score 20.4; DB 3; Length 25;  
Best Local Similarity 75.0%; Pred. No. 1.9;  
Matches 18; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGATCTAAGTTG 24  
Db 1 GTTCAGCTTTTCTACWAATG 24

RESULT 8  
US-08-894-511-7  
Sequence 7, Application US/08894511  
Patent No. 614330  
GENERAL INFORMATION:  
APPLICANT: CROUZET, Joel  
APPLICANT: SCHERMAN, Daniel  
APPLICANT: CAMERON, Beatrice  
APPLICANT: WILS, Pierre  
APPLICANT: DARQUET, Anne-Marie  
TITLE OF INVENTION: DNA MOLECULES, PREPARATION THEREOF AND  
USE THEREOF IN GENE THERAPY  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, Mailstop 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,511  
FILING DATE: 21-FEB-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02117  
FILING DATE: 23-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO FR96/00274  
FILING DATE: 21-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Savitzky Esq., Martin F.  
REGISTRATION NUMBER: 29,699  
REFERENCE/DOCKET NUMBER: ST95013-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3808  
TELEFAX: (610) 454-3816  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Oligonucleotide"

US-08-894-511-7

Query Match 80.8%; Score 19.4; DB 3; Length 21;  
Best Local Similarity 95.2%; Pred. No. 5.3;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CAGCTTTTGATCTAAGTTG 24  
Db 1 CAGCTTTTCTAAGTTG 21

RESULT 9  
US-09-655-728-7  
Sequence 7, Application US/09655728  
Patent No. 6492164  
GENERAL INFORMATION:  
APPLICANT: CROUZET, Joel  
APPLICANT: SCHERMAN, Daniel  
APPLICANT: CAMERON, Beatrice  
APPLICANT: WILS, Pierre  
APPLICANT: DARQUET, Anne-Marie  
TITLE OF INVENTION: DNA MOLECULES, PREPARATION THEREOF AND  
USE THEREOF IN GENE THERAPY  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, Mailstop 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/655,728  
FILING DATE: 05-Sep-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,511  
FILING DATE: <Unknown>  
APPLICATION NUMBER: WO FR96/00274  
FILING DATE: 21-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Savitzky Esq., Martin F.  
REGISTRATION NUMBER: 29,699  
REFERENCE/DOCKET NUMBER: ST95013-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3816  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Oligonucleotide"

US-09-655-728-7

Query Match 80.8%; Score 19.4; DB 4; Length 21;  
Best Local Similarity 95.2%; Pred. No. 5.3;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CAGCTTTTGATCTAAGTTG 24  
Db 1 CAGCTTTTCTAAGTTG 21

RESULT 10  
US-09-233-493-15  
Sequence 15, Application US/09233493  
Patent No. 6143557  
GENERAL INFORMATION:  
APPLICANT: Hartley, James L.  
APPLICANT: Brasch, Michael A.  
TITLE OF INVENTION: Recombinational Cloning Using Engineered  
NUMBER OF SEQUENCES: 35

RESULT 9  
US-09-655-728-7  
Sequence 7, Application US/09655728  
Patent No. 6492164  
GENERAL INFORMATION:  
APPLICANT: CROUZET, Joel  
APPLICANT: SCHERMAN, Daniel  
APPLICANT: CAMERON, Beatrice  
APPLICANT: WILS, Pierre  
APPLICANT: DARQUET, Anne-Marie  
TITLE OF INVENTION: DNA MOLECULES, PREPARATION THEREOF AND  
USE THEREOF IN GENE THERAPY  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, Mailstop 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/655,728  
FILING DATE: 05-Sep-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/894,511  
FILING DATE: <Unknown>  
APPLICATION NUMBER: WO FR96/00274  
FILING DATE: 21-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Savitzky Esq., Martin F.  
REGISTRATION NUMBER: 29,699  
REFERENCE/DOCKET NUMBER: ST95013-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3816  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Oligonucleotide"

US-09-655-728-7

Query Match 80.8%; Score 19.4; DB 4; Length 21;  
Best Local Similarity 95.2%; Pred. No. 5.3;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CAGCTTTTGATCTAAGTTG 24  
Db 1 CAGCTTTTCTAAGTTG 21

RESULT 10  
US-09-233-493-15  
Sequence 15, Application US/09233493  
Patent No. 6143557  
GENERAL INFORMATION:  
APPLICANT: Hartley, James L.  
APPLICANT: Brasch, Michael A.  
TITLE OF INVENTION: Recombinational Cloning Using Engineered  
NUMBER OF SEQUENCES: 35

US-09-233-493-15

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
;; STREET: 1100 New York Ave., N. W. Suite 600  
;; CITY: Washington  
;; STATE: DC  
;; COUNTRY: USA  
;; ZIP: 20005-3934  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; FILING DATE: 07-JUN-1996  
;; APPLICATION NUMBER: 08/663,002  
;; FILING DATE: 07-JUN-1996  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 09/005,476  
;; FILING DATE: 12-JAN-1998  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/663,002  
;; FILING DATE: 07-JUN-1996  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/486,139  
;; FILING DATE: 07-JUN-1995  
;; CLASSIFICATION:  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-2600  
;; TELEFAX: 202-371-2540  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 25 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: both  
;; TOPOLOGY: both  
;; MOLECULE TYPE: cdna  
;; US-09-233-493-15

Query Match 80.0%; Score 19.2; DB 3; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||  
Db 1 GTTCAGCTTTTGTACTAAGTTG 24

RESULT 11  
US-09-005-476-15  
; Sequence 15, Application US/09005476  
; Patent No. 6171861  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; TITLE OF INVENTION: Recombinational Cloning Using Engineered  
; TITLE OF INVENTION: Recombination Sites  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
; STREET: 1100 New York Ave., N. W. Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/005,476

;; FILING DATE: herewith  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/663,002  
;; FILING DATE: 07-JUN-1996  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-2600  
;; TELEFAX: 202-371-2540  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 25 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: both  
;; TOPOLOGY: both  
;; MOLECULE TYPE: cdna  
;; US-09-005-476-15

Query Match 80.0%; Score 19.2; DB 3; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||  
Db 1 GTTCAGCTTTTGTACTAAGTTG 24

RESULT 12  
US-09-233-492-15  
; Sequence 15, Application US/09233492  
; Patent No. 6270969  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; TITLE OF INVENTION: Recombinational Cloning Using Engineered  
; TITLE OF INVENTION: Recombination Sites  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
; STREET: 1100 New York Ave., N. W. Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/233,492  
; FILING DATE: 20-JAN-1999  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/663,002  
; FILING DATE: 07-JUN-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/486,139  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cdna  
; US-09-233-492-15

Query Match

80.0%; Score 19.2; DB 3; Length 25;

Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATACACTAGTTG 24  
|||||  
Db 1 GTTCAGCTTTTGATACAAAGTTG 24

RESULT 13  
US-09-296-280-15  
; Sequence 15, Application US/09296280  
; Patent No. 6277608  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; APPLICANT: Temple, Gary F.  
; APPLICANT: Fox, Donna K.  
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having  
; TITLE OF INVENTION: Recombination Sites  
; FILE REFERENCE: 0942.2850007  
; CURRENT APPLICATION NUMBER: US/09/296,280  
; CURRENT FILING DATE: 1999-04-22  
; EARLIER APPLICATION NUMBER: US 09/177,387  
; EARLIER FILING DATE: 1998-10-23  
; EARLIER APPLICATION NUMBER: US 60/065,930  
; EARLIER FILING DATE: 1997-10-24  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patentin ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: recombination  
; OTHER INFORMATION: products  
US-09-296-280-15

Query Match 80.0%; Score 19.2; DB 3; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATACACTAGTTG 24  
|||||  
Db 1 GTTCAGCTTTTGATACAAAGTTG 24

RESULT 14  
US-09-498-074-15  
; Sequence 15, Application US/09498074  
; Patent No. 6534264  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; TITLE OF INVENTION: Recombinational Cloning Using Engineered  
; TITLE OF INVENTION: Recombination Sites  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C  
; STREET: 1100 New York Ave., N. W. Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/498,074  
; FILING DATE: (Herewith)  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/005,476  
; FILING DATE: 12-JAN-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/663,002  
; FILING DATE: 07-JUN-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/486,139  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cdna  
US-09-498-074-15

Query Match 80.0%; Score 19.2; DB 4; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.5;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATACACTAGTTG 24  
|||||  
Db 1 GTTCAGCTTTTGATACAAAGTTG 24

RESULT 15  
PCT-US96-10082A-15  
; Sequence 15, Application PC/TUS9610082A  
; GENERAL INFORMATION:  
; APPLICANT: Life Technologies, Inc.  
; APPLICANT: 8717 Grovemont Circle  
; APPLICANT: Gaithersburg, MD 20884-9980  
; APPLICANT: United States of America  
; APPLICANT: Brasch, Michael A.  
; TITLE OF INVENTION: Recombinational Cloning Using Engineered  
; TITLE OF INVENTION: Recombination Sites  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C  
; STREET: 1100 New York Ave., N. W. Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10082A  
; FILING DATE: 07-JUN-1996  
; CLASSIFICATION:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cdna  
PCT-US96-10082A-15

Query Match 80.0%; Score 19.2; DB 5; Length 25;

Best Local Similarity 87.5%; Pred. No. 6.5;  
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 GTTCAGCTTTTGTGATACTAAGTTG 24  
 |||||  
 Db 1 GTTCAGCTTTTGTGACAAAGTTG 24  
 |||||

Search completed: December 3, 2003, 14:51:46  
 Job time : 29.562 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:41 ; Search time 114.843 Seconds  
(without alignments)  
564.131 Million cell updates/sec

Title: US-10-082-772-3

Perfect score: 24

Sequence: 1 gttcagcttttgataactagtgtg 24

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*

2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*

3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*

4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*

5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*

6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*

7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*

8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*

9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*

10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*

11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*

12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*

13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*

14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*

15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*

16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*

17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*

18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*

19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*

20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*

21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*

22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*

23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*

24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	100.0	24	AAF61394	Expression vector
2	24	100.0	24	AAF61395	Expression vector
3	24	100.0	27	AAF61425	AttP* DNA fragment
4	24	100.0	243	AAF61421	E. coli attP* DNA
5	22.4	93.3	27	AAF61175	Phage-lambda recom
6	22.4	93.3	27	AAF61424	AttP DNA fragment
7	22.4	93.3	27	ABZ58728	Att site nucleotid
8	22.4	93.3	201	AAQ50319	Detection probe 11

9	22.4	93.3	243	22	AAF61418	Bacteriophage lamb
10	22.4	93.3	282	25	ACC44676	Murine rDNA PCR pr
11	22.4	93.3	282	25	AET15636	Artificial plant c
12	22.4	93.3	610	22	AAF79770	Bacteriophage lamb
13	22.4	93.3	1763	20	AAZ30709	Rat neuronal immed
14	22.4	93.3	4346	25	ACC44716	Plasmid pSV40191at
15	22.4	93.3	4346	25	AET15615	Artificial plant c
16	22.4	93.3	4909	16	AAF18924	Plasmid pAI261. S
17	22.4	93.3	5641	22	AAF30800	Vector plasmid pJM
18	22.4	93.3	5670	22	AAF30801	Vector plasmid pJM
19	22.4	93.3	5826	22	AAF30798	Vector plasmid pDY
20	22.4	93.3	6071	22	AAF30799	Vector plasmid pDY
21	22.4	93.3	7852	14	AAQ45682	Sequencing vector
22	21.2	88.3	25	20	AAZ78977	Oligonucleotide #4
23	20.8	86.7	27	22	AAZ506177	Phage-lambda recom
24	20.8	86.7	27	25	ABZ58730	Att site nucleotid
25	20.8	86.7	162	22	AAF61420	E. coli attR DNA f
26	20.4	85.0	25	20	AAZ78976	Oligonucleotide #4
27	19.4	80.8	21	17	AAZ73771	Bacteriophage lamb
28	19.4	80.8	21	25	ACC44644	Nucleotide core re
29	19.4	80.8	48	21	AAZ55575	Mutational attL si
30	19.4	80.8	48	22	AAZ506251	PCR primer attL3G
31	19.2	80.0	25	18	AAZ48224	attP1 core region.
32	19.2	80.0	25	20	AAZ78949	Oligonucleotide #1
33	19.2	80.0	25	22	AAZ14443	Recombination site
34	19.2	80.0	25	22	AAZ506205	Phage-lambda recom
35	19.2	80.0	25	22	AAZ506209	Recombination site
36	19.2	80.0	25	22	AAZ55749	Recombination site
37	19.2	80.0	25	22	AAZ87880	Escherichia coli c
38	19.2	80.0	25	24	AAQ82127	Core sequence of r
39	19.2	80.0	25	25	ACC44664	Recombination site
40	19.2	80.0	25	25	ABZ58758	Att site nucleotid
41	19.2	80.0	25	25	ABZ58762	Att site nucleotid
42	19.2	80.0	25	25	AET15634	Artificial plant c
43	19.2	80.0	27	22	AAZ506179	Phage-lambda recom
44	19.2	80.0	27	22	AAZ506203	Phage-lambda recom
45	19.2	80.0	27	22	AAZ506207	Phage-lambda recom

#### ALIGNMENTS

#### RESULT 1

AAF61394 ID AAF61394 standard; DNA; 24 BP.

XX AC AAF61394;

XX DT 05-JUN-2001 (first entry)

XX DE Expression vector pXEX-2-XR attP and attP\* PCR primer O3.

XX DE Sequence-specific recombination; SSR; integrase; gene therapy; somatic; targeted integration; attP; attP\*; PCR primer; ss.

XX OS Synthetic.

XX PN DE19941186-A1.

XX PD 01-MAR-2001.

XX PF 30-AUG-1999; 99DE-1041186.

XX PR 30-AUG-1999; 99DE-1041186.

XX PA (DROE/) DROEGE P.

XX PI Droeg P;

XX DR WPI; 2001-246016/26.

XX PT Sequence-specific recombination of DNA in eukaryotes, useful particularly for somatic cell gene therapy, uses an integrase to effect



PT recombination between att sites -  
XX Example 1; Page 8; 24pp; German.  
XX  
CC This invention describes a novel sequence-specific recombination (SSR) of  
CC DNA in a eukaryotic cell, comprising introducing two DNA sequences (I,  
CC II) into a cell, using an integrase (int) to effect SSR. The invention  
CC also describes (1) a nucleic acid comprising a 243 base pair sequence  
CC (III), fully defined in the specification, or its derivatives; and (2)  
CC vector containing (III), or its derivatives, plus a therapeutic gene, or  
CC its derivatives. The method is particularly used in somatic gene therapy  
CC in humans and animals, but can be applied more generally for gene  
CC transfer to animal or plant cells. The method is simple and controllable,  
CC and provides stable and targeted integration of selected DNA sequences.  
XX  
SQ Sequence 24 BP; 5 A; 3 C; 5 G; 11 T; 0 other;  
  
Query Match 100.0%; Score 24; DB 22; Length 24;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GTTCAGCTTTTGATCTAAGTTG 24  
DB 1 GTTCAGCTTTTGATCTAAGTTG 24  
  
RESULT 2  
AAF61395/c  
ID AAF61395 standard; DNA; 24 BP.  
XX  
AC AAF61395;  
XX  
DT 05-JUN-2001 (first entry)  
XX  
DE Expression vector pXEX-2-XR attP and attP\* PCR primer O4.  
XX  
KW Sequence-specific recombination; SSR; integrase; gene therapy; somatic;  
KW targeted integration; attP; PCR primer; ss.  
XX  
OS Synthetic.  
XX  
XX DE19941186-A1.  
XX  
PD 01-MAR-2001.  
XX  
PF 30-AUG-1999; 99DE-1041186.  
XX  
PR 30-AUG-1999; 99DE-1041186.  
XX  
PA (DROE/) DROEGE P.  
XX  
PI Droege P;  
XX  
DR WPI; 2001-246016/26.  
XX  
PT Sequence-specific recombination of DNA in eukaryotes, useful  
PT particularly for somatic cell gene therapy, uses an integrase to effect  
PT recombination between att sites -  
XX  
PS Disclosure; Fig 6B; 24pp; German.  
XX  
CC This invention describes a novel sequence-specific recombination (SSR) of  
CC DNA in a eukaryotic cell, comprising introducing two DNA sequences (I,  
CC II) into a cell, using an integrase (int) to effect SSR. The invention  
CC also describes (1) a nucleic acid comprising a 243 base pair sequence  
CC (III), fully defined in the specification, or its derivatives; and (2)  
CC vector containing (III), or its derivatives, plus a therapeutic gene, or  
CC its derivatives. The method is particularly used in somatic gene therapy  
CC in humans and animals, but can be applied more generally for gene  
CC transfer to animal or plant cells. The method is simple and controllable,  
CC and provides stable and targeted integration of selected DNA sequences.  
XX  
SQ Sequence 24 BP; 11 A; 5 C; 3 G; 5 T; 0 other;  
  
Query Match 100.0%; Score 24; DB 22; Length 24;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GTTCAGCTTTTGATCTAAGTTG 24  
DB 1 GTTCAGCTTTTGATCTAAGTTG 24  
  
RESULT 3  
AAF61425  
ID AAF61425 standard; DNA; 27 BP.  
XX  
AC AAF61425;  
XX  
DT 05-JUN-2001 (first entry)  
XX  
DE AttP\* DNA fragment.  
XX  
KW Sequence-specific recombination; SSR; integrase; gene therapy; somatic;  
KW targeted integration; attP; ds.  
XX  
OS Unidentified.  
XX  
XX DE19941186-A1.  
XX  
PD 01-MAR-2001.  
XX  
PF 30-AUG-1999; 99DE-1041186.  
XX  
PR 30-AUG-1999; 99DE-1041186.  
XX  
PA (DROE/) DROEGE P.  
XX  
PI Droege P;  
XX  
DR WPI; 2001-246016/26.  
XX  
PT Sequence-specific recombination of DNA in eukaryotes, useful  
PT particularly for somatic cell gene therapy, uses an integrase to effect  
PT recombination between att sites -  
XX  
PS Disclosure; Fig 6B; 24pp; German.  
XX  
CC This invention describes a novel sequence-specific recombination (SSR) of  
CC DNA in a eukaryotic cell, comprising introducing two DNA sequences (I,  
CC II) into a cell, using an integrase (int) to effect SSR. The invention  
CC also describes (1) a nucleic acid comprising a 243 base pair sequence  
CC (III), fully defined in the specification, or its derivatives; and (2)  
CC vector containing (III), or its derivatives, plus a therapeutic gene, or  
CC its derivatives. The method is particularly used in somatic gene therapy  
CC in humans and animals, but can be applied more generally for gene  
CC transfer to animal or plant cells. The method is simple and controllable,  
CC and provides stable and targeted integration of selected DNA sequences.  
XX  
SQ Sequence 27 BP; 6 A; 4 C; 6 G; 11 T; 0 other;  
  
Query Match 100.0%; Score 24; DB 22; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GTTCAGCTTTTGATCTAAGTTG 24  
DB 1 GTTCAGCTTTTGATCTAAGTTG 24  
  
RESULT 4  
AAF61421  
ID AAF61421 standard; DNA; 243 BP.  
XX  
AC AAF61421;  
XX

DT 05-JUN-2001 (first entry)  
XX E. coli attP\* DNA fragment.  
DE Sequence-specific recombination; SSR; integrase; gene therapy; somatic;  
KW targeted integration; attP; ds.  
XX Escherichia coli.  
OS DE19941186-A1.  
XX 01-MAR-2001.  
XX 30-AUG-1999; 99DE-1041186.  
XX 30-AUG-1999; 99DE-1041186.  
XX (DROE/) DROEGE P.  
XX Droege P;  
XX WPI; 2001-246016/26.  
XX Sequence-specific recombination of DNA in eukaryotes, useful  
PT particularly for somatic cell gene therapy, uses an integrase to effect  
PT recombination between att sites -  
XX Claim 19; Page 14; 24pp; German.  
XX This invention describes a novel sequence-specific recombination (SSR) of  
CC DNA in a eukaryotic cell, comprising introducing two DNA sequences (I,  
CC II) into a cell, using an integrase (Int) to effect SSR. The invention  
CC also describes (1) a nucleic acid comprising a 243 base pair sequence  
CC (III), fully defined in the specification, or its derivatives; and (2)  
CC vector containing (III), or its derivatives, plus a therapeutic gene, or  
CC its derivatives. The method is particularly used in somatic gene therapy  
CC in humans and animals, but can be applied more generally for gene  
CC transfer to animal or plant cells. The method is simple and controllable,  
CC and provides stable and targeted integration of selected DNA sequences.  
XX Sequence 243 BP; 74 A; 34 C; 34 G; 101 T; 0 other;  
SQ Query Match 100.0%; Score 24; DB 22; Length 243;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
Db 139 GTTCAGCTTTTGTACTAAGTTG 162  
RESULT 5  
AAS06175  
ID AAS06175 standard; DNA; 27 BP.  
XX AAS06175;  
AC 12-SEP-2001 (first entry)  
XX Phage-lambda recombination site atp01.  
DE Bacteriophage lambda; recombination; att site; PCR primer; lambda Int;  
KW lambda integrase; therapeutic; ss.  
XX Bacteriophage lambda.  
OS WO200142509-A1.  
XX 14-JUN-2001.  
XX 11-DEC-2000; 2000WO-US33546.  
XX 10-DEC-1999; 99US-0169983.

PR 09-MAR-2000; 2000US-0188020.  
XX (CHEO/) CHEO D.  
PA (BRAS/) BRASCH M A.  
PA (TEMP/) TEMPLE G F.  
PA (HART/) HARTLEY J L.  
PA (BYRD/) BYRD D R N.  
XX Cheo D, Brasch MA, Temple GF, Hartley JL, Byrd DRN;  
XX WPI; 2001-356174/37.  
XX Producing hybrid nucleic acids, useful for expressing novel therapeutic  
PT polypeptides, by mixing the same or different nucleic acids having one  
PT or more recombination sites in the presence of recombination proteins,  
XX e.g. Cre -  
XX Disclosure; Fig 24A; 357pp; English.  
XX AAS06174-AAS06322 represent Bacteriophage lambda att recombination  
CC site nucleic acid sequences, and PCR primers of the invention. The  
CC att sequences are recognised by the recombination protein lambda  
CC integrase (Int). The invention is a new method of producing a population  
CC of hybrid nucleic acids comprising mixing at least a first population of  
CC nucleic acids comprising one or more recombination sites with at least  
CC one target nucleic acid comprising one or more recombination sites and  
CC causing some or all of the nucleic acids to recombine with all or some of  
CC the target nucleic acids. The method is useful for producing a population  
CC of hybrid nucleic acids which may be the same or different. The nucleic  
CC acids may be used to express therapeutic proteins or peptides and they  
CC may also be used to create novel fusion proteins by expressing different  
CC sequences linked to each other. The method allows simultaneous cloning of  
XX two or more different nucleic acids.  
SQ Sequence 27 BP; 6 A; 4 C; 5 G; 12 T; 0 other;  
Query Match 93.3%; Score 22.4; DB 22; Length 27;  
Best Local Similarity 95.8%; Pred. No. 1.4;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
Db 1 GTTCAGCTTTTGTACTAAGTTG 24  
RESULT 6  
AAS061424  
ID AAF61424 standard; DNA; 27 BP.  
XX AAF61424;  
AC 05-JUN-2001 (first entry)  
XX AttP DNA fragment.  
DE Sequence-specific recombination; SSR; integrase; gene therapy; somatic;  
KW targeted integration; attP; ds.  
XX Unidentified.  
OS DE19941186-A1.  
XX 01-MAR-2001.  
XX 30-AUG-1999; 99DE-1041186.  
XX 30-AUG-1999; 99DE-1041186.  
XX (DROE/) DROEGE P.  
XX Droege P;  
XX WPI; 2001-246016/26.

XX Sequence-specific recombination of DNA in eukaryotes, useful  
PT particularly for somatic cell gene therapy, uses an integrase to effect  
PT recombination between att sites -  
XX  
XX Disclosure; Fig 6B; 24pp; German.  
XX  
XX This invention describes a novel sequence-specific recombination (SSR) of  
CC DNA in a eukaryotic cell, comprising introducing two DNA sequences (I,  
CC II) into a cell, using an integrase (Int) to effect SSR. The invention  
CC also describes (1) a nucleic acid comprising a 243 base pair sequence  
CC (III), fully defined in the specification, or its derivatives; and (2)  
CC vector containing (III), or its derivatives, plus a therapeutic gene, or  
CC its derivatives. The method is particularly used in somatic gene therapy  
CC in humans and animals, but can be applied more generally for gene  
CC transfer to animal or plant cells. The method is simple and controllable,  
CC and provides stable and targeted integration of selected DNA sequences.  
XX  
XX Sequence 27 BP; 6 A; 4 C; 5 G; 12 T; 0 other;  
SQ  
Query Match 93.3%; Score 22.4; DB 22; Length 27;  
Best Local Similarity 95.8%; Pred. No. 1.4;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGGATCTAAGTTG 24  
DB 1 GTTCAGCTTTTGGATCTAAGTTG 24  
RESULT 7  
ID ABZ58728 standard; DNA; 27 BP.  
XX  
XX ABZ58728;  
DT 01-MAY-2003 (first entry)  
DE Att site nucleotide sequence attP0.  
KW Nucleic acid insertion; recombination; nucleic acid selection;  
KW nucleic acid isolation; att; ds.  
XX Synthetic.  
XX WO200295055-A2.  
PN 28-NOV-2002.  
XX  
XX 21-MAY-2002; 2002WO-US15947.  
PF  
XX 21-MAY-2001; 2001US-291973P.  
PR  
XX (INVI-) INVITROGEN CORP.  
PA  
XX Brasch MA, Cheo D, Li X, Espesito D, Byrd DRN;  
PI  
XX WPI; 2003-129436/12.  
DR  
XX Inserting a population of nucleic acids into a second target molecule  
PT for selecting and isolating nucleic acid molecules by mixing the second  
PT population of nucleic acid with a second target nucleic acid -  
XX  
XX Disclosure; Fig 13A; 273pp; English.  
XX  
XX The invention relates to inserting a population of nucleic acids into a  
CC second target molecule. The method involves (a) mixing a first population  
CC of nucleic acid comprising one or more recombination sites with a target  
CC nucleic acid; (b) causing some or all of the nucleic acid molecules of  
CC the first population to recombine with the first target nucleic acid  
CC molecules to form a second population; (c) mixing the second population  
CC of nucleic acid with a second target nucleic acid; and (d) causing some  
CC or all of the nucleic acid molecules of the second population to  
CC recombine with some or all of the second target nucleic acid molecules to

CC form a third population of nucleic acid. The method is useful for  
CC selecting and isolating nucleic acid molecules. Sequences ABZ58727-762  
CC represent att recombination site sequences used in the method of the  
CC invention.  
XX  
XX Sequence 27 BP; 6 A; 4 C; 5 G; 12 T; 0 other;  
SQ  
Query Match 93.3%; Score 22.4; DB 25; Length 27;  
Best Local Similarity 95.8%; Pred. No. 1.4;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGGATCTAAGTTG 24  
DB 1 GTTCAGCTTTTGGATCTAAGTTG 24  
RESULT 8  
ID AAQ50319 standard; DNA; 201 BP.  
XX  
XX AAQ50319;  
AC  
XX 27-APR-1994 (first entry)  
DT  
XX Detection probe 11.  
DE  
XX KW Probe; chip; cell; detection; label; capture probe; ss.  
XX Synthetic.  
XX JP05236997-A.  
PN  
XX 17-SEP-1993.  
PD  
XX 28-FEB-1992; 92JP-0042829.  
PF  
XX 28-FEB-1992; 92JP-0042829.  
PR  
XX (HITA) HITACHI LTD.  
PA  
XX WPI; 1993-330600/42.  
DR  
XX Chip for capturing polynucleotide - has several different  
PT complementary probes fixed on cells at different sites on single  
PT tip  
XX  
XX Disclosure; Page 8; 10pp; Japanese.  
XX  
XX The sequences given in AAQ50309-26 are probes which were used on a  
CC chip for capturing a number of target polynucleotides. A group of  
CC these probes may be attached on cells which form part of a single  
CC chip. The oligonucleotides to be detected are labelled and are  
CC complementary to the set of capture probes used. The detection chip  
CC may be used to detect a number of different polynucleotides  
CC concurrently.  
XX  
XX Sequence 201 BP; 58 A; 38 C; 27 G; 78 T; 0 other;  
SQ  
Query Match 93.3%; Score 22.4; DB 14; Length 201;  
Best Local Similarity 95.8%; Pred. No. 1.6;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGGATCTAAGTTG 24  
DB 40 GTTCAGCTTTTGGATCTAAGTTG 63  
RESULT 9  
ID AAF61418 standard; DNA; 243 BP.  
XX  
XX AAF61418;  
AC  
XX

DT 05-JUN-2001 (first entry)  
XX Bacteriophage lambda attP DNA fragment.  
DE Sequence-specific recombination; SSR; integrase; gene therapy; somatic;  
XX targeted integration; attP; ds.  
KW Bacteriophage lambda.  
OS DE19941186-A1.  
XX PN 01-MAR-2001.  
PD 30-AUG-1999; 99DE-1041186.  
XX PF 30-AUG-1999; 99DE-1041186.  
XX PR (DROE/) DROEGE P.  
XX PA Droeg P;  
XX PI WPI; 2001-246016/26.  
XX DR Sequence-specific recombination of DNA in eukaryotes, useful  
PT particularly for somatic cell gene therapy, uses an integrase to effect  
PT recombination between att sites -  
XX  
PS Claim 2; Page 13; 24pp; German.  
XX  
CC This invention describes a novel sequence-specific recombination (SSR) of  
CC DNA in a eukaryotic cell, comprising introducing two DNA sequences (I,  
CC II) into a cell, using an integrase (Int) to effect SSR. The invention  
CC also describes (1) a nucleic acid comprising a 243 base pair sequence  
CC (III), fully defined in the specification, or its derivatives; and (2)  
CC vector containing (III), or its derivatives, plus a therapeutic gene, or  
CC its derivatives. The method is particularly used in somatic gene therapy  
CC in humans and animals, but can be applied more generally for gene  
CC transfer to animal or plant cells. The method is simple and controllable,  
CC and provides stable and targeted integration of selected DNA sequences.  
XX  
SQ Sequence 243 BP; 74 A; 34 C; 33 G; 102 T; 0 other;  
Query Match 93.3%; Score 22.4; DB 22; Length 243;  
Best Local Similarity 95.8%; Pred. No. 1.7;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
Db 139 GTTCAGCTTTTGTACTAAGTTG 162  
RESULT 10  
ACC44676  
ID ACC44676 standard; DNA; 282 BP.  
XX  
XX ACC44676;  
AC  
XX  
XX 29-MAY-2003 (first entry)  
DT  
XX Murine rDNA PCR primer SEQ ID NO:72.  
DE  
XX Chromosome-based platform; artificial chromosome; eukaryotic chromosome;  
KW att site; integrase; recombinase; ACes; gene therapy; transgenic animal;  
KW platform artificial chromosome expression system; gene; ds.  
XX  
XX Mus musculus.  
OS Synthetic.  
OS  
XX WO200297059-A2.  
XX PN 05-DEC-2002.  
XX PD 30-MAY-2002; 2002WO-US17452.  
XX PF

XX 30-MAY-2001; 2001US-294758P.  
PR 21-MAR-2002; 2002US-366891P.  
XX  
XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.  
XX  
XX Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E;  
PI Stewart S, Shellard J;  
XX WPI; 2003-140461/13.  
DR  
XX Novel eukaryotic chromosome comprising one or many att sites which  
XX permits site-directed integration in the presence of lambda-integrase,  
PT useful for site-specific recombination-directed integration of DNA of  
PT interest -  
XX  
XX Example 2; Page 207; 272pp; English.  
XX  
CC The present invention describes a eukaryotic chromosome (I) comprising  
CC one or several att sites, where an att site is heterologous to the  
CC chromosome, and permits site-directed integration in the presence of  
CC lambda-integrase. Also described: (1) a platform artificial chromosome  
CC expression system (ACes) (II) comprising several sites that participate  
CC in recombinase catalysed recombination; and (2) a method (MI) for  
CC introducing a heterologous nucleic acid into a platform artificial  
CC chromosome. (I) can be used in gene therapy. (MI) is useful for  
CC introducing a heterologous nucleic acid molecule into a platform  
CC artificial chromosome, preferably an ACes. (II) is useful for producing a  
CC transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or  
CC mammal) by introducing (II) by cell fusion, lipid-mediated transfection  
CC by a carrier system, microinjection, microcell fusion, electroporation,  
CC microprojectile bombardment or direct DNA transfer into an embryonic  
CC cell, preferably a stem cell or an embryo. (II) comprises a heterologous  
CC nucleic acid that encodes a therapeutic product which is useful for  
CC making a library of ACes comprising random portions of a genome. ACC44612  
CC to ACC44732 and ABP96650 to ABP96657 represent sequences used in the  
CC exemplification of the present invention.  
XX  
XX Sequence 282 BP; 79 A; 49 C; 39 G; 115 T; 0 other;  
SQ  
Query Match 93.3%; Score 22.4; DB 25; Length 282;  
Best Local Similarity 95.8%; Pred. No. 1.7;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
Db 153 GTTCAGCTTTTGTACTAAGTTG 176  
RESULT 11  
ABT16636  
ID ABT16636 standard; DNA; 282 BP.  
XX  
XX ABT16636;  
AC  
XX  
XX 03-APR-2003 (first entry)  
DT  
XX Artificial plant chromosome related oligo SEQ ID No 48.  
DE  
XX Plant artificial chromosome; PAC; transgenic plant; vaccine;  
KW blood factor; herbicide; stress; agronomic; nutrient quality;  
KW bacterial artificial chromosome; BAC; yeast artificial chromosome; YAC;  
KW ds.  
XX  
XX Unidentified.  
OS  
XX WO200296923-A1.  
XX PN 05-DEC-2002.  
XX PD 30-MAY-2002; 2002WO-US17451.  
XX PF 30-MAY-2001; 2001US-294687P.  
XX PR



transduction. The identification of neuronal IEGs and the proteins they encode may provide important information about the function of neurons in, for example, learning, memory, synaptic transmission, tolerance and neuronal plasticity. Neuronal IEGs, neuronal IEG protein products, cells expressing neuronal IEGs and antibodies against neuronal IEG proteins can be used to treat an animal with a deficiency in neuronal IEG responsiveness to stimuli, such that the effect of the deficiency is minimised. The deficiency may be a reduced or elevated level of expression of an IEG. The neuronal stimulus comprises a maximal electroconvulsive seizure and its effects influence learning or memory. The IEGs and protein products are useful in identifying compounds that modulate the expression or activity of IEG nucleic acids or proteins, respectively.

Query Match 93.3%; Score 22.4; DB 20; Length 1763;  
Best Local Similarity 95.8%; Pred. No. 1.9;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||  
Db 293 GTTCAGCTTTTGTACTAAGTTG 270

RESULT 14  
ID ACC44716  
XX ACC44716 standard; DNA; 4346 BP.  
AC ACC44716;  
XX  
DT 29-MAY-2003 (first entry)  
DE Plasmid pSV40193attPeensePUR nucleotide sequence SEQ ID NO:113.  
XX  
KW Chromosome-based platform; artificial chromosome; eukaryotic chromosome;  
KW att site; integrase; recombinase; ACes; gene therapy; transgenic animal;  
KW platform artificial chromosome expression system; gene; ds.  
XX  
OS Bacteriophage lambda.  
OS Synthetic.  
XX  
FN WO200297059-A2.  
XX  
PD 05-DEC-2002.  
XX  
PF 30-MAY-2002; 2002WO-US17452.  
XX  
PR 30-MAY-2001; 2001US-294758P.  
PR 21-MAR-2002; 2002US-366891P.  
XX  
PA (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.

PI Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E;  
PI Stewart S, Sheppard J;  
XX  
DR WPI; 2003-140461/13.  
XX  
PT Novel eukaryotic chromosome comprising one or many att sites which  
PT permits site-directed integration in the presence of lambda-integrase,  
PT useful for site-specific recombination-directed integration of DNA of  
PT interest -  
XX  
PS Example 3; Page 244-245; 272pp; English.

The present invention describes a eukaryotic chromosome (I) comprising one or several att sites, where an att site is heterologous to the chromosome, and permits site-directed integration in the presence of lambda-integrase. Also described: (1) a platform artificial chromosome expression system (ACes) (II) comprising several sites that participate in recombination catalysed recombination; and (2) a method (M1) for introducing a heterologous nucleic acid into a platform artificial chromosome. (I) can be used in gene therapy. (M1) is useful for

introducing a heterologous nucleic acid molecule into a platform artificial chromosome, preferably an ACes. (II) is useful for producing a transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or mammal) by introducing (II) by cell fusion, lipid-mediated transfection, by a carrier system, microinjection, microcell fusion, electroporation, microprojectile bombardment or direct DNA transfer into an embryonic cell, preferably a stem cell or an embryo. (II) comprises a heterologous nucleic acid that encodes a therapeutic product which is useful for making a library of ACes comprising random portions of a genome. ACC44612 to ACC44732 and AB996650 to AB996657 represent sequences used in the exemplification of the present invention.

Query Match 93.3%; Score 22.4; DB 25; Length 4346;  
Best Local Similarity 95.8%; Pred. No. 2;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||  
Db 4208 GTTCAGCTTTTGTACTAAGTTG 4231

RESULT 15  
ABT16615  
XX ABT16615 standard; DNA; 4346 BP.  
AC ABT16615;  
XX  
DT 03-APR-2003 (first entry)  
DE Artificial plant chromosome related plasmid DNA SEQ ID No 26.  
XX  
KW Plant artificial chromosome; PAC; transgenic plant; vaccine;  
KW blood factor; herbicide; stress; agronomical; nutrient quality;  
KW Bacterial artificial chromosome; BAC; yeast artificial chromosome; YAC;  
KW ds.  
XX  
OS Unidentified.  
XX  
FN WO200296923-A1.  
XX  
PD 05-DEC-2002.  
XX  
PF 30-MAY-2002; 2002WO-US17451.  
XX  
PR 30-MAY-2001; 2001US-294687P.  
PR 04-JUN-2001; 2001US-296329P.  
XX  
PA (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.

PI Perez C, Fabijanski SF, Perkins E;  
PI WPI; 2003-140436/13.  
XX  
PT Producing artificial chromosome by introducing a nucleic acid into  
PT plant cell, selecting artificial chromosome that has one or more repeat  
PT regions with equivalent amounts of euchromatic and heterochromatic  
PT nucleic acids -  
XX  
PS Example 19; Page 255-256; 269pp; English.

The invention relates to a novel method for producing plant artificial chromosomes. The invention also relates to methods for targeting insertion of heterologous DNA into plant artificial chromosomes, methods for delivery of plant chromosomes to selected cells and tissues. The isolated plant artificial chromosome (PAC) is useful for producing a transgenic plant, which involves introducing the PAC into a plant cell. The PAC comprises a heterologous nucleic acid encoding a gene product such as enzymes, antisense RNA, rDNA, structural proteins, marker proteins, ligands, receptors, ribozymes, therapeutic proteins, and biopharmaceutical proteins, vaccines, blood factors, antigens, hormones,

CC cytokines, growth factors, antibodies, or a product that provides for  
CC resistance to diseases, insects, herbicides, or stress in a plant. The  
CC heterologous nucleic acid optionally encodes a product that provides an  
CC agronomically important trait in the plant, e.g. a product that alters  
CC nutrient use and/or improves the nutrient quality of the plant. The  
CC heterologous nucleic acid is contained within a bacterial artificial  
CC chromosome (BAC) or a yeast artificial chromosome (YAC). This  
CC polynucleotide sequence represents the DNA of a plasmid used in the  
CC method of the invention.

XX  
SQ Sequence 4346 BP; 1022 A; 1208 C; 1094 G; 1022 T; 0 other;

Query Match 93.3%; Score 22.4; DB 25; Length 4346;  
Best Local Similarity 95.8%; Pred. No. 2;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTGACTACTAGTTG 24  
|||  
Db 4208 GTTCAGCTTTTGTGACTACTAGTTG 4231

Search completed: December 3, 2003, 12:23:35  
Job time : 114.843 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 13:18:51 ; Search time 107.107 Seconds  
(without alignments)  
744.732 Million cell updates/sec

Title: US-10-082-772-4

Perfect score: 24

Sequence: 1 caacttagtattcaaaaagctgaac 24

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
C 1	24	100.0	24	14	US-10-082-772-3
C 2	24	100.0	24	14	US-10-082-772-4
C 3	22.4	93.3	27	9	US-09-732-914-2
C 4	22.4	93.3	27	14	US-10-151-690-26
C 5	22.4	93.3	282	14	US-10-161-403-72
C 6	22.4	93.3	1763	12	US-09-244-805-57
C 7	22.4	93.3	1763	12	US-09-245-277-57
C 8	22.4	93.3	4346	14	US-10-161-403-113
C 9	21.2	88.3	25	9	US-09-855-797A-43
C 10	21.2	88.3	25	10	US-09-907-900-43
C 11	21.2	88.3	25	10	US-09-907-719-43
C 12	21.2	88.3	25	12	US-09-985-448-43
C 13	21.2	88.3	25	12	US-10-300-892-43
C 14	20.8	86.7	27	9	US-09-732-914-4
C 15	20.8	86.7	27	14	US-10-151-690-28
C 16	20.4	85.0	25	9	US-09-855-797A-42

C 17	20.4	85.0	25	10	US-09-907-900-42	Sequence 42, Appl
C 18	20.4	85.0	25	10	US-09-907-719-42	Sequence 42, Appl
C 19	20.4	85.0	25	12	US-09-985-448-42	Sequence 42, Appl
C 20	20.4	85.0	25	12	US-10-300-892-42	Sequence 42, Appl
C 21	19.4	80.8	21	11	US-09-981-803-4	Sequence 4, Appl
C 22	19.4	80.8	21	14	US-10-252-384-7	Sequence 7, Appl
C 23	19.4	80.8	21	14	US-10-161-403-33	Sequence 33, Appl
C 24	19.4	80.8	48	9	US-09-732-914-70	Sequence 70, Appl
C 25	19.2	80.0	25	9	US-09-732-914-36	Sequence 32, Appl
C 26	19.2	80.0	25	9	US-09-732-914-36	Sequence 36, Appl
C 27	19.2	80.0	25	9	US-09-855-797A-15	Sequence 15, Appl
C 28	19.2	80.0	25	10	US-09-907-900-15	Sequence 15, Appl
C 29	19.2	80.0	25	10	US-09-907-719-15	Sequence 15, Appl
C 30	19.2	80.0	25	11	US-09-432-085-15	Sequence 15, Appl
C 31	19.2	80.0	25	12	US-09-985-448-15	Sequence 15, Appl
C 32	19.2	80.0	25	12	US-10-300-892-15	Sequence 15, Appl
C 33	19.2	80.0	25	14	US-10-055-001A-10	Sequence 10, Appl
C 34	19.2	80.0	25	14	US-10-058-292-15	Sequence 15, Appl
C 35	19.2	80.0	25	14	US-10-162-879-15	Sequence 15, Appl
C 36	19.2	80.0	25	14	US-10-161-403-55	Sequence 55, Appl
C 37	19.2	80.0	25	14	US-10-151-690-56	Sequence 56, Appl
C 38	19.2	80.0	25	14	US-10-151-690-60	Sequence 60, Appl
C 39	19.2	80.0	27	9	US-09-732-914-6	Sequence 6, Appl
C 40	19.2	80.0	27	9	US-09-732-914-30	Sequence 30, Appl
C 41	19.2	80.0	27	9	US-09-732-914-34	Sequence 34, Appl
C 42	19.2	80.0	27	14	US-10-151-690-30	Sequence 30, Appl
C 43	19.2	80.0	27	14	US-10-151-690-54	Sequence 54, Appl
C 44	19.2	80.0	27	14	US-10-151-690-58	Sequence 58, Appl
C 45	19.2	80.0	4470	14	US-10-151-690-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-10-082-772-3/c  
; Sequence 3, Application US/10082772  
; Publication No. US20030027337A1  
; GENERAL INFORMATION:  
; APPLICANT: DROGE, PETER  
; APPLICANT: CHRIST, NICOLE  
; APPLICANT: LOREBACH, ELKE  
; TITLE OF INVENTION: SEQUENCE-SPECIFIC DNA RECOMBINATION IN EUKARYOTIC CELLS  
; FILE REFERENCE: DEBE:008US  
; CURRENT APPLICATION NUMBER: US/10/082,772  
; CURRENT FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: PCT/DE 00/02947  
; PRIOR FILING DATE: 2000-08-29  
; PRIOR APPLICATION NUMBER: DE 199 41 186.7  
; PRIOR FILING DATE: 1999-08-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Primer  
US-10-082-772-3

Query Match 100.0%; Score 24; DB 14; Length 24;  
Best Local Similarity 100.0%; Pred. No. 0.31; Mismatches 0; Indels 0; Gaps 0;  
Matches 24; Conservative 0

Qy 1 CAACCTAGTATCAAAAAGCTGAAC 24

Db 24 CAACCTAGTATCAAAAAGCTGAAC 1

RESULT 2

US-10-082-772-4  
; Sequence 4, Application US/10082772



Publication No. US20030027337A1  
; GENERAL INFORMATION:  
; APPLICANT: DROGE, PETER  
; APPLICANT: CHRIST, NICOLE  
; APPLICANT: LORBACH, ELKE  
; TITLE OF INVENTION: SEQUENCE-SPECIFIC DNA RECOMBINATION IN EUKARYOTIC CELLS  
; FILE REFERENCE: DEBE:008US  
; CURRENT APPLICATION NUMBER: US/10/082,772  
; CURRENT FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: PCT/DE 00/02947  
; PRIOR FILING DATE: 2000-08-29  
; PRIOR APPLICATION NUMBER: DE 199 41 186.7  
; PRIOR FILING DATE: 1999-08-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-082-772-4  
Query Match 100.0%; Score 24; DB 14; Length 24;  
Best Local Similarity 100.0%; Pred. No. 0.31; Mismatches 0; Indels 0; Gaps 0;  
Matches 24; Conservative 0;  
QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
Db 1 CAACCTAGTATCAAAAAGCTGAAC 24  
RESULT 3  
US-09-732-914-2/c  
; Sequence 2, Application US/09732914  
; Patent No. US20020007051A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheo, David  
; APPLICANT: Brasch, Michael A.  
; APPLICANT: Temple, Gary F.  
; APPLICANT: Hartley, James L.  
; APPLICANT: Byrd, Devon R.N.  
; TITLE OF INVENTION: Use of Multiple Recombination Sites with Unique Specificity in  
; FILE REFERENCE: 0942.5010002  
; CURRENT APPLICATION NUMBER: US/09/732,914  
; CURRENT FILING DATE: 2000-12-11  
; PRIOR APPLICATION NUMBER: US 60/169,983  
; PRIOR FILING DATE: 1999-12-10  
; PRIOR APPLICATION NUMBER: US 60/188,020  
; PRIOR FILING DATE: 2000-03-09  
; NUMBER OF SEQ ID NOS: 140  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 27  
; TYPE: DNA  
; ORGANISM: attP0  
US-09-732-914-2  
Query Match 93.3%; Score 22.4; DB 9; Length 27;  
Best Local Similarity 95.8%; Pred. No. 1.6;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
Db 24 CAACCTAGTATCAAAAAGCTGAAC 1  
RESULT 4  
US-10-151-690-26/c  
; Sequence 26, Application US/10151690  
; Publication No. US2003012455A1

GENERAL INFORMATION:  
; APPLICANT: BRASCH, MICHAEL A.  
; APPLICANT: CHEO, DAVID  
; APPLICANT: LI, XIAO  
; APPLICANT: ESPOSITO, DOMINIC  
; APPLICANT: BYRD, DEVON R.N.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR USE IN ISOLATION OF NUCLEIC ACID MOLE  
; FILE REFERENCE: 0942.5120001  
; CURRENT APPLICATION NUMBER: US/10/151,690  
; CURRENT FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: US 10/151,690  
; PRIOR FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: US 60/291,973  
; PRIOR FILING DATE: 2001-05-21  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 27  
; TYPE: DNA  
; ORGANISM: attP0  
US-10-151-690-26  
Query Match 93.3%; Score 22.4; DB 14; Length 27;  
Best Local Similarity 95.8%; Pred. No. 1.6;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
Db 24 CAACCTAGTATCAAAAAGCTGAAC 1  
RESULT 5  
US-10-161-403-72/c  
; Sequence 72, Application US/10161403  
; Publication No. US20030119104A1  
; GENERAL INFORMATION:  
; APPLICANT: Perkins, Edward  
; APPLICANT: Perez, Carl  
; APPLICANT: Lindenbaum, Michael  
; APPLICANT: Greene, Amy  
; APPLICANT: Leung, Josephine  
; APPLICANT: Fleming, Elena  
; APPLICANT: Stewart, Sandra  
; APPLICANT: Shellard, Joan  
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
; FILE REFERENCE: 24601-420  
; CURRENT APPLICATION NUMBER: US/10/161,403  
; CURRENT FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: 60/294,758  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/366,891  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 72  
; LENGTH: 282  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: attp  
US-10-161-403-72  
Query Match 93.3%; Score 22.4; DB 14; Length 282;  
Best Local Similarity 95.8%; Pred. No. 2.6;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
Db 176 CAACCTAGTATCAAAAAGCTGAAC 153  
RESULT 6  
US-09-244-805-57

; Sequence 57, Application US/09244805  
; Publication No. US20030203840A1  
; GENERAL INFORMATION:  
; APPLICANT: Worley, Paul F.  
; APPLICANT: Lanahan, Anthony  
; APPLICANT: Goetz, Bernard  
; APPLICANT: Heimisch, Holger  
; APPLICANT: Kuner, Rohini  
; APPLICANT: Scheek, Sigrid  
; APPLICANT: Nikolich, Karoly  
; APPLICANT: Zhukovski, Eugene  
; TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: 10496/004001  
; CURRENT APPLICATION NUMBER: US/09/244,805  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/074,518  
; EARLIER FILING DATE: 1998-02-12  
; EARLIER APPLICATION NUMBER: 60/074,135  
; EARLIER FILING DATE: 1998-02-06  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 57  
; LENGTH: 1763  
; TYPE: DNA  
; ORGANISM: Eukaryote  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(1763)  
; OTHER INFORMATION: Y = C or T  
; OTHER INFORMATION: n = A,T,C or G  
US-09-244-805-57

Query Match 93.3%; Score 22.4; DB 12; Length 1763;  
Best Local Similarity 95.8%; Pred. No. 3.7;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACTTAGTATCAAAAAGCTGAAC 24  
|||  
Db 270 CAACTTAGTATCAAAAAGCTGAAC 293

RESULT 7  
US-09-245-277-57  
; Sequence 57, Application US/09245277  
; Publication No. US20030211984A1  
; GENERAL INFORMATION:  
; APPLICANT: Worley, Paul F.  
; APPLICANT: Lanahan, Anthony  
; TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: JHU1530-3  
; CURRENT APPLICATION NUMBER: US/09/245,277  
; CURRENT FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/074,518  
; PRIOR FILING DATE: 1998-02-12  
; PRIOR APPLICATION NUMBER: 60/074,135  
; PRIOR FILING DATE: 1998-02-06  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 57  
; LENGTH: 1763  
; TYPE: DNA  
; ORGANISM: Eukaryote  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(1763)  
; OTHER INFORMATION: Y = C or T  
; OTHER INFORMATION: n = A,T,C or G  
US-09-245-277-57

Query Match 93.3%; Score 22.4; DB 12; Length 1763;  
Best Local Similarity 95.8%; Pred. No. 3.7;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CAACTTAGTATCAAAAAGCTGAAC 24  
|||  
Db 270 CAACTTAGTATCAAAAAGCTGAAC 293

RESULT 8  
US-10-161-403-113/c  
; Sequence 113, Application US/10161403  
; Publication No. US20030119104A1  
; GENERAL INFORMATION:  
; APPLICANT: Perkins, Edward  
; APPLICANT: Perez, Carl  
; APPLICANT: Lindenbaum, Michael  
; APPLICANT: Greene, Amy  
; APPLICANT: Leung, Josephine  
; APPLICANT: Fleming, Elena  
; APPLICANT: Stewart, Sandra  
; APPLICANT: Shellard, Joan  
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
; FILE REFERENCE: 24601-420  
; CURRENT APPLICATION NUMBER: US/10/161,403  
; CURRENT FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: 60/294,758  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/366,891  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 113  
; LENGTH: 4346  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pSV40-193AttpsensePur Plasmid  
US-10-161-403-113

Query Match 93.3%; Score 22.4; DB 14; Length 4346;  
Best Local Similarity 95.8%; Pred. No. 4.4;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACTTAGTATCAAAAAGCTGAAC 24  
|||  
Db 4231 CAACTTAGTATCAAAAAGCTGAAC 4208

RESULT 9  
US-09-855-797A-43/c  
; Sequence 43, Application US/09855797A  
; Patent No. US20020094574A1  
; GENERAL INFORMATION:  
; APPLICANT: Hartley, James L.  
; APPLICANT: Brasch, Michael A.  
; APPLICANT: Temple, Gary F.  
; APPLICANT: Fox, Donna K.  
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having  
; FILE REFERENCE: 0942.285008  
; CURRENT APPLICATION NUMBER: US/09/855,797A  
; CURRENT FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 09/296,281  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: US 60/065,930  
; PRIOR FILING DATE: 1997-10-24  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 43  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: recombination

OTHER INFORMATION: products  
US-09-855-797A-43

Query Match 88.3%; Score 21.2; DB 9; Length 25;  
Best Local Similarity 83.3%; Pred. No. 5.6;  
Matches 20; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
||| : ||| : ||| :  
Db 24 CAACCTTWGTAYAAAAAGCTGAAC 1

RESULT 10

```

US-09-907-900-43/c
; Sequence 43, Application US/09907900
; Patent No. US20020172997A1
; GENERAL INFORMATION:
; APPLICANT: Hartley, James L.
; APPLICANT: Brasch, Michael A.
; APPLICANT: Temple, Gary F.
; APPLICANT: Fox, Donna K.
; TITLE OF INVENTION: Recombination Cloning Using Nucleic Acids Having
; TITLE OF INVENTION: Recombination Sites
; FILE REFERENCE: 0942.2850004
; CURRENT APPLICATION NUMBER: US/09/907,900
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 09/177,387
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 43
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: recombination
; OTHER INFORMATION: products
US-09-907-900-43

```

Query Match 88.3%; Score 21.2; DB 10; Length 25;  
Best Local Similarity 83.3%; Pred. No. 5.6;  
Matches 20; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY            1 CAACTTAGTATCAAAAAGCTGAAC 24  
             |||||:||||:|||||  
Db            24 CAACTTGTAYAAAAAGCTGAAC 1

## RESULT 11

```

US-09-907-719-43/c
/ Sequence 43, Application US/09907719
/ Publication No. US20020192819A1
/ GENERAL INFORMATION:
/ APPLICANT: Hartley, James L.
/ APPLICANT: Brasch, Michael A.
/ APPLICANT: Temple, Gary F.
/ APPLICANT: Fox, Donna K.
/ TITLE OF INVENTION: Recombination Cloning Using Nucleic Acids Having
/ FILE REFERENCE: 0942.2850004
/ CURRENT APPLICATION NUMBER: US/09/907,719
/ CURRENT FILING DATE: 2001-07-19
/ PRIOR APPLICATION NUMBER: US/09/177,387
/ PRIOR FILING DATE: 1998-10-23
/ NUMBER OF SEQ ID NOS: 60
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 43
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Unknown
/ FEATURES:
/ OTHER INFORMATION: Description of Unknown Organism: recombination
/ OTHER INFORMATION: products

```

US-09-907-719-43

Query Match	88.3%	Score 21.2;	DB 10;	Length 25;
Best Local Similarity	83.3%	Pred. No. 5.6;		
Matches 20; Conservative	3;	Mismatches 1;	Indels 0;	Gaps 0;

**QY**            1 CAAC<sup>T</sup>TAGTATCAAAAAGCTGAAC 24  
               |||||:||||:|:|||||  
**Db**            24 CAAC<sup>T</sup>TGWTAYAAAAAGCTGAAC 1

RESULT 12

```

US-09-985-448-43/C
; Sequence #3, Application US/09985448
; Publication No. US20030157716A1
; GENERAL INFORMATION:
; APPLICANT: Hartley, James L.
; APPLICANT: Brasch, Michael A.
; APPLICANT: Temple, Gary F.
; APPLICANT: Fox, Donna K.
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having
; TITLE OF INVENTION: Recombination Sites
; FILE REFERENCE: 0942-2850004
; CURRENT APPLICATION NUMBER: US/09/985,448
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US/09/177,387
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: US 60/065,930
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 43
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Unknown
; FEATURES:
; OTHER INFORMATION: Description of Unknown Organism: recombination
US-09-985-448-43

```

Query Match 88.3%; Score 21.2; DB 12; Length 25;  
Best Local Similarity 83.3%; Pred. No. 5.6;  
Matches 20; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy            1 CAACCTTAGTATCAAAAAGCTGAAC 24  
             |||||:||||:|||||  
Db            24 CAACTTGTATYAAPAAAGCTGAAC 1

RESULT 13

```

US-10-300-892-43/c
; Sequence 43, Application US/10300892
; Publication No. US20030175970A1
; GENERAL INFORMATION:
; APPLICANT: Hartley, James L.
; APPLICANT: Brasch, Michael A.
; APPLICANT: Temple, Gary F.
; APPLICANT: Fox, Donna K.
; TITLE OF INVENTION: Recombination Cloning Using Nucleic Acids Having
; FILE REFERENCE: 0942.2850004
; CURRENT APPLICATION NUMBER: US/10/300,892
; CURRENT FILING DATE: 2002-11-21
; PRIORITY APPLICATION NUMBER: US/09/907,719
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US/09/177,387
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Unknown

```

```
;
;
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: recombination
; OTHER INFORMATION: products
US-10-300-892-43

Query Match      88.3%; Score 21.2; DB 12; Length 25;
Best Local Similarity 83.3%; Pred. No. 5.6;
Matches 20; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 CAACTTAGTATCAAAAAGCTGAAC 24
      |||||:||||:|||||
Db      24 CAACTTGTATAYAAAAAGCTGAAC 1

RESULT 14
US-09-732-914-4/c
; Sequence 4, Application US/09732914
; Patent No. US20020007051A1
; GENERAL INFORMATION:
; APPLICANT: Cheo, David
; APPLICANT: Brasch, Michael A.
; APPLICANT: Temple, Gary F.
; APPLICANT: Hartley, James L.
; APPLICANT: Byrd, Devon R.N.
; TITLE OF INVENTION: Use of Multiple Recombination Sites with Unique Specificity in
; FILE OF INVENTION: Recombinational Cloning
; FILE REFERENCE: 0942.5010002
; CURRENT APPLICATION NUMBER: US/09/732,914
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/169,993
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: US 60/188,020
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 27
; TYPE: DNA
; ORGANISM: attRO
US-09-732-914-4

Query Match      86.7%; Score 20.8; DB 9; Length 27;
Best Local Similarity 91.7%; Pred. No. 8.5;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CAACTTAGTATCAAAAAGCTGAAC 24
      |||||:|||||
Db      24 CAACTTAGTATCAAAAAGCTGAAC 1

RESULT 15
US-10-151-690-28/c
; Sequence 28, Application US/10151690
; Publication No. US20030124555A1
; GENERAL INFORMATION:
; APPLICANT: BRASCH, MICHAEL A.
; APPLICANT: CHEO, DAVID
; APPLICANT: LI, XIAO
; APPLICANT: ESPOSITO, DOMINIC
; APPLICANT: BYRD, DEVON R.N.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR USE IN ISOLATION OF NUCLEIC ACID MOI
; FILE REFERENCE: 0942.5120001
; CURRENT APPLICATION NUMBER: US/10/151,690
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 10/151,690
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/291,973
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 27
; TYPE: DNA
```

```
; ORGANISM: attRO
US-10-151-690-28

Query Match      86.7%; Score 20.8; DB 14; Length 27;
Best Local Similarity 91.7%; Pred. No. 8.5;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CAACTTAGTATCAAAAAGCTGAAC 24
      |||||:|||||
Db      24 CAACTTAGTATCAAAAAGCTGAAC 1

Search completed: December 3, 2003, 17:22:25
Job time : 107.107 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 12:13:41 ; Search time 114.843 Seconds  
(without alignments)  
564.131 Million cell updates/sec

Title: US-10-082-772-4

Perfect score: 24

Sequence: 1 caacttagatcaaaaagctgaac 24

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03:

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	24	100.0	24	AAF61394	Expression vector
C 2	24	100.0	24	AAF61395	Expression vector
C 3	24	100.0	27	AAF61425	AttP* DNA fragment
C 4	24	100.0	243	AAF61421	E. coli attP* DNA
C 5	22.4	93.3	27	AAF61421	E. coli attP* DNA
C 6	22.4	93.3	27	AAF61421	Phage-lambda recom
C 7	22.4	93.3	27	AAF61424	AttP DNA fragment
C 8	22.4	93.3	201	AAQ50319	Detection probe 11

C 9	22.4	93.3	243	22	AAF61418	Bacteriophage lamb
C 10	22.4	93.3	282	25	ACC44676	Murine rDNA PCR pr
C 11	22.4	93.3	282	25	ABT16836	Artificial plant c
C 12	22.4	93.3	610	22	AAF79770	Bacteriophage lamb
C 13	22.4	93.3	1763	20	AAZ30709	Rat neuronal immed
C 14	22.4	93.3	4346	25	ACC44716	Plasmid pSV40193at
C 15	22.4	93.3	4346	25	ABT16615	Artificial plant c
C 16	22.4	93.3	4909	16	AAI18924	Plasmid pAI261. S
C 17	22.4	93.3	5641	22	AAF30800	Vector plasmid pUM
C 18	22.4	93.3	5670	22	AAF30801	Vector plasmid pUM
C 19	22.4	93.3	5826	22	AAF30798	Vector plasmid pBY
C 20	22.4	93.3	6071	22	AAF30799	Vector plasmid pBY
C 21	22.4	93.3	7652	14	AAQ45682	Sequencing vector
C 22	21.2	88.3	25	20	AAK78977	Oligonucleotide #4
C 23	20.8	86.7	27	22	AAQ506177	Phage-lambda recom
C 24	20.8	86.7	27	25	ABZ58730	Att site nucleotid
C 25	20.8	86.7	162	22	AAF61420	E. coli attR DNA f
C 26	20.4	85.0	25	20	AAK78976	Oligonucleotide #4
C 27	19.4	80.8	21	17	AAI37371	Bacteriophage lamb
C 28	19.4	80.8	21	25	ACC44644	Nucleotide core re
C 29	19.4	80.8	48	21	AAQ55575	Mutational attL si
C 30	19.4	80.8	48	22	AAQ56251	PCR primer attL3G
C 31	19.2	80.0	25	18	AAI48224	attP1 core region.
C 32	19.2	80.0	25	20	AAK78949	Oligonucleotide #1
C 33	19.2	80.0	25	22	AAI14443	Recombination site
C 34	19.2	80.0	25	22	AAQ56209	Phage-lambda recom
C 35	19.2	80.0	25	22	AAQ56205	Recombination site
C 36	19.2	80.0	25	22	AAQ55749	Escherichia coli c
C 37	19.2	80.0	25	22	AAQ87880	Core sequence of r
C 38	19.2	80.0	25	24	AAQ82127	Recombination site
C 39	19.2	80.0	25	25	ACC44664	Att site nucleotid
C 40	19.2	80.0	25	25	ABZ58758	Att site nucleotid
C 41	19.2	80.0	25	25	ABZ58762	Artificial plant c
C 42	19.2	80.0	25	25	ABT16634	Phage-lambda recom
C 43	19.2	80.0	27	22	AAQ506179	Phage-lambda recom
C 44	19.2	80.0	27	22	AAQ506203	Phage-lambda recom
C 45	19.2	80.0	27	22	AAQ506207	Phage-lambda recom

ALIGNMENTS

RESULT 1

AAF61394/c

ID AAF61394 standard; DNA; 24 Bp.

AC AAF61394;

DT 05-JUN-2001 (first entry)

DE Expression vector pKEX-2-XR attP and attP\* PCR primer O3.

DE Sequence-specific recombination; SSR; integrase; gene therapy; somatic;

DE targeted integration; attP; attP\*; PCR primer; ss.

OS Synthetic.

PN DE19941186-A1.

PD 01-MAR-2001.

PF 30-AUG-1999; 99DE-1041186.

PR 30-AUG-1999; 99DE-1041186.

XX (DROE/) DROEGE P.

XX Droege P;

XX WPI; 2001-246016/26.

XX Sequence-specific recombination of DNA in eukaryotes, useful

XX particularly for somatic cell gene therapy, uses an integrase to effect



DT 05-JUN-2001 (first entry)  
XX E. coli attP\* DNA fragment.  
DE Sequence-specific recombination; SSR; integrase; gene therapy; somatic;  
XX targeted integration; attP; ds.  
KW Escherichia coli.  
XX DE19941186-A1.  
XX 01-MAR-2001.  
XX 30-AUG-1999; 99DE-1041186.  
XX 30-AUG-1999; 99DE-1041186.  
XX (DROE/) DROEGE P.  
XX Droege P;  
XX WPI; 2001-246016/26.  
XX  
XX Sequence-specific recombination of DNA in eukaryotes, useful  
PT particularly for somatic cell gene therapy, uses an integrase to effect  
PT recombination between att sites -  
XX  
XX Claim 19; Page 14; 24pp; German.  
XX  
XX This invention describes a novel sequence-specific recombination (SSR) of  
CC DNA in a eukaryotic cell, comprising introducing two DNA sequences (I,  
CC II) into a cell, using an integrase (Int) to effect SSR. The invention  
CC also describes (1) a nucleic acid comprising a 243 base pair sequence  
CC (III), fully defined in the specification, or its derivatives; and (2)  
CC vector containing (III), or its derivatives, plus a therapeutic gene, or  
CC its derivatives. The method is particularly used in somatic gene therapy  
CC in humans and animals, but can be applied more generally for gene  
CC transfer to animal or plant cells. The method is simple and controllable,  
CC and provides stable and targeted integration of selected DNA sequences.  
XX  
XX Sequence 243 BP; 74 A; 34 C; 34 G; 101 T; 0 other;  
SQ  
Query Match 100.0%; Score 24; DB 22; Length 243;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
DB 162 CAACCTAGTATCAAAAAGCTGAAC 139  
RESULT 5  
AAS06175/c  
ID AAS06175 standard; DNA; 27 BP.  
XX  
XX AAS06175;  
XX 12-SEP-2001 (first entry)  
XX Phage-lambda recombination site atP01.  
XX Bacteriophage lambda; recombination; att site; PCR primer; lambda Int;  
XX lambda integrase; therapeutic; ss.  
XX Bacteriophage lambda.  
XX WO200142509-A1.  
XX 14-JUN-2001.  
XX 11-DEC-2000; 2000WO-US33546.  
XX 10-DEC-1999; 99US-0169983.

PR 09-MAR-2000; 2000US-0188020.  
XX (CHEO/) CHEO D.  
XX (BRAS/) BRASCH M A.  
XX (TEMP/) TEMPLE G F.  
XX (HART/) HARTLEY J L.  
XX (BYRD/) BYRD D R N.  
XX  
XX Cheo D, Brasch MA, Temple GF, Hartley JL, Byrd DRN;  
XX WPI; 2001-356174/37.  
XX  
XX Producing hybrid nucleic acids, useful for expressing novel therapeutic  
PT polypeptides, by mixing the same or different nucleic acids having one  
PT or more recombination sites in the presence of recombination proteins,  
XX e.g. Cre -  
XX  
XX Disclosure; Fig 24A; 357pp; English.  
XX  
XX AAS06174-AAS06322 represent Bacteriophage lambda att recombination  
CC site nucleic acid sequences, and PCR primers of the invention. The  
CC att sequences are recognised by the recombination protein lambda  
CC integrase (Int). The invention is a new method of producing a population  
CC of hybrid nucleic acids comprising mixing at least a first population of  
CC nucleic acids comprising one or more recombination sites with at least  
CC one target nucleic acid comprising one or more recombination sites and  
CC causing some or all of the nucleic acids to recombine with all or some of  
CC the target nucleic acids. The method is useful for producing a population  
CC of hybrid nucleic acids which may be the same or different. The nucleic  
CC acids may be used to express therapeutic proteins or peptides and they  
CC may also be used to create novel fusion proteins by expressing different  
CC sequences linked to each other. The method allows simultaneous cloning of  
XX two or more different nucleic acids.  
XX  
XX Sequence 27 BP; 6 A; 4 C; 5 G; 12 T; 0 other;  
SQ  
Query Match 93.3%; Score 22.4; DB 22; Length 27;  
Best Local Similarity 95.8%; Pred. No. 1.4;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
DB 24 CAACCTAGTATCAAAAAGCTGAAC 1  
RESULT 6  
AAF61424/c  
ID AAF61424 standard; DNA; 27 BP.  
XX  
XX AAF61424;  
XX  
XX 05-JUN-2001 (first entry)  
XX  
XX AttP DNA fragment.  
XX  
XX Sequence-specific recombination; SSR; integrase; gene therapy; somatic;  
KW targeted integration; attP; ds.  
XX Unidentified.  
XX  
XX DE19941186-A1.  
XX  
XX 01-MAR-2001.  
XX  
XX 30-AUG-1999; 99DE-1041186.  
XX  
XX 30-AUG-1999; 99DE-1041186.  
XX (DROE/) DROEGE P.  
XX Droege P;  
XX WPI; 2001-246016/26.

XX Sequence-specific recombination of DNA in eukaryotes, useful  
 PT particularly for somatic cell gene therapy, uses an integrase to effect  
 PT recombination between att sites -  
 XX  
 XX Disclosure; Fig 6B; 24pp; German.

XX This invention describes a novel sequence-specific recombination (SSR) of  
 CC DNA in a eukaryotic cell, comprising introducing two DNA sequences (I,  
 CC II) into a cell, using an integrase (Int) to effect SSR. The invention  
 CC also describes (1) a nucleic acid comprising a 243 base pair sequence  
 CC (III), fully defined in the specification, or its derivatives; and (2)  
 CC vector containing (III), or its derivatives, plus a therapeutic gene, or  
 CC its derivatives. The method is particularly used in somatic gene therapy  
 CC in humans and animals, but can be applied more generally for gene  
 CC transfer to animal or plant cells. The method is simple and controllable,  
 CC and provides stable and targeted integration of selected DNA sequences.

XX Sequence 27 BP; 6 A; 4 C; 5 G; 12 T; 0 other;

Query Match 93.3%; Score 22.4; DB 22; Length 27;  
 Best Local Similarity 95.8%; Pred. No. 1.4;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
 DB 24 CAACCTTAGTATCAAAAAGCTGAAC 1

RESULT 7  
 ABZ58728/c  
 ID ABZ58728 standard; DNA; 27 BP.  
 XX  
 AC ABZ58728;  
 XX  
 XX 01-MAY-2003 (first entry)  
 XX  
 DE Att site nucleotide sequence attP0.  
 XX  
 XX Nucleic acid insertion; recombination; nucleic acid selection;  
 KW nucleic acid isolation; att; ds.  
 XX  
 OS Synthetic.

XX WO200295055-A2.

XX 28-NOV-2002.

XX 21-MAY-2002; 2002WO-US15947.

XX 21-MAY-2001; 2001US-291973P.

XX (INVI-) INVITROGEN CORP.

XX Brasch MA, Cheo D, Li X, Esposito D, Byrd DRN;

XX WPI; 2003-129436/12.

XX Inserting a population of nucleic acids into a second target molecule  
 PT for selecting and isolating nucleic acid molecules by mixing the second  
 PT population of nucleic acid with a second target nucleic acid -

XX Disclosure; Fig 13A; 273pp; English.

XX The invention relates to inserting a population of nucleic acids into a  
 CC second target molecule. The method involves (a) mixing a first population  
 CC of nucleic acid comprising one or more recombination sites with a target  
 CC nucleic acid; (b) causing some or all of the nucleic acid molecules of  
 CC the first population to recombine with the first target nucleic acid  
 CC molecules to form a second population; (c) mixing the second population  
 CC of nucleic acid with a second target nucleic acid; and (d) causing some  
 CC or all of the nucleic acid molecules of the second population to  
 CC recombine with some or all of the second target nucleic acid molecules to

CC form a third population of nucleic acid. The method is useful for  
 CC selecting and isolating nucleic acid molecules. Sequences ABZ58727-762  
 CC represent att recombination site sequences used in the method of the  
 CC invention.

XX Sequence 27 BP; 6 A; 4 C; 5 G; 12 T; 0 other;

Query Match 93.3%; Score 22.4; DB 25; Length 27;  
 Best Local Similarity 95.8%; Pred. No. 1.4;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
 DB 24 CAACCTTAGTATCAAAAAGCTGAAC 1

RESULT 8  
 AAQ50319/c  
 ID AAQ50319 standard; DNA; 201 BP.  
 XX  
 AC AAQ50319;  
 XX  
 DT 27-APR-1994 (first entry)  
 XX  
 DE Detection probe 11.  
 XX  
 KW Probe; chip; cell; detection; label; capture probe; ss.  
 XX  
 OS Synthetic.

XX JP05236997-A.

XX 17-SEP-1993.

XX 28-FEB-1992; 92JP-0042829.

XX 28-FEB-1992; 92JP-0042829.

XX (HITA) HITACHI LTD.

XX WPI; 1993-330600/42.

XX Chip for capturing polynucleotide - has several different  
 PT complementary probes fixed on cells at different sites on single  
 PT tip

XX Disclosure; Page 8; 10pp; Japanese.

XX The sequences given in AAQ50309-26 are probes which were used on a  
 CC chip for capturing a number of target polynucleotides. A group of  
 CC these probes may be attached on cells which form part of a single  
 CC chip. The oligonucleotides to be detected are labelled and are  
 CC complementary to the set of capture probes used. The detection chip  
 CC may be used to detect a number of different polynucleotides  
 CC concurrently.

XX Sequence 201 BP; 58 A; 38 C; 27 G; 78 T; 0 other;

Query Match 93.3%; Score 22.4; DB 14; Length 201;  
 Best Local Similarity 95.8%; Pred. No. 1.6;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTTAGTATCAAAAAGCTGAAC 24  
 DB 63 CAACCTTAGTATCAAAAAGCTGAAC 40

RESULT 9  
 AAF61418/c  
 ID AAF61418 standard; DNA; 243 BP.  
 XX  
 XX AAF61418;  
 AC  
 XX



DT 05-JUN-2001 (first entry)  
XX Bacteriophage lambda attP DNA fragment.  
DE Sequence-specific recombination; SSR; integrase; gene therapy; somatic;  
XX targeted integration; attP; ds.  
KW Bacteriophage lambda.  
XX DE19941186-A1.  
XX 01-MAR-2001.  
PD 30-AUG-1999; 99DE-1041186.  
XX 30-AUG-1999; 99DE-1041186.  
XX (DROE/) DROEGE P.  
PA Droege P;  
XX WPI; 2001-246016/26.  
DR Sequence-specific recombination of DNA in eukaryotes, useful  
XX particularly for somatic cell gene therapy, uses an integrase to effect  
PT recombination between att sites -  
XX Claim 2; Page 13; 24pp; German.  
PS This invention describes a novel sequence-specific recombination (SSR) of  
XX DNA in a eukaryotic cell, comprising introducing two DNA sequences (I, II,  
CC I) into a cell, using an integrase (Int) to effect SSR. The invention  
CC also describes (1) a nucleic acid comprising a 243 base pair sequence  
CC (III), fully defined in the specification, or its derivatives; and (2)  
CC vector containing (III), or its derivatives, plus a therapeutic gene, or  
CC its derivatives. The method is particularly used in somatic gene therapy  
CC in humans and animals, but can be applied more generally for gene  
CC transfer to animal or plant cells. The method is simple and controllable,  
CC and provides stable and targeted integration of selected DNA sequences.  
XX SQ Sequence 243 BP; 74 A; 34 C; 33 G; 102 T; 0 other;  
Query Match 93.3%; Score 22.4; DB 22; Length 243;  
Best Local Similarity 95.8%; Pred. No. 1.7;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
Db 162 CAACCTAGTATCAAAAAGCTGAAC 139  
RESULT 10  
ACCA4676/c  
ID ACCA4676 standard; DNA; 282 BP.  
XX ACCA4676;  
AC ACCA4676;  
XX 29-MAY-2003 (first entry)  
DE Murine rDNA PCR primer SEQ ID NO:72.  
XX Chromosome-based platform; artificial chromosome; eukaryotic chromosome;  
KW att site; integrase; recombination; ACes; gene therapy; transgenic animal;  
XX platform artificial chromosome expression system; gene; ds.  
XX Mus musculus.  
OS Synthetic.  
OS WO200297059-A2.  
XX 05-DEC-2002.  
PD 30-MAY-2002; 2002WO-US17452.  
XX 30-MAY-2002; 2002WO-US17452.  
PF

XX 30-MAY-2001; 2001US-294758P.  
PR 21-MAR-2002; 2002US-366891P.  
XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.  
PA Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E;  
XX Stewart S, Shellard J;  
PI WPI; 2003-140461/13.  
XX Novel eukaryotic chromosome comprising one or many att sites which  
XX permits site-directed integration in the presence of lambda-integrase,  
PT useful for site-specific recombination-directed integration of DNA of  
PT interest -  
XX Example 2; Page 207; 272pp; English.  
PS The present invention describes a eukaryotic chromosome (I) comprising  
XX one or several att sites, where an att site is heterologous to the  
CC chromosome, and permits site-directed integration in the presence of  
CC lambda-integrase. Also described: (1) a platform artificial chromosome  
CC expression system (ACes) (II) comprising several sites that participate  
CC in recombination catalysed recombination; and (2) a method (M1) for  
CC introducing a heterologous nucleic acid into a platform artificial  
CC chromosome. (1) can be used in gene therapy. (M1) is useful for  
CC introducing a heterologous nucleic acid molecule into a platform  
CC artificial chromosome, preferably an ACes. (II) is useful for producing a  
CC transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or  
CC mammal) by introducing (II) by cell fusion, lipid-mediated transfection,  
CC microprojectile bombardment or direct DNA transfer into an embryonic  
CC cell, preferably a stem cell or an embryo. (II) comprises a heterologous  
CC nucleic acid that encodes a therapeutic product which is useful for  
CC making a library of ACes comprising random portions of a genome. ACC44612  
CC to ACC44732 and ABP96650 to ABP96657 represent sequences used in the  
CC exemplification of the present invention.  
XX SQ Sequence 282 BP; 79 A; 49 C; 39 G; 115 T; 0 other;  
Query Match 93.3%; Score 22.4; DB 25; Length 282;  
Best Local Similarity 95.8%; Pred. No. 1.7;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
Db 176 CAACCTAGTATCAAAAAGCTGAAC 153  
RESULT 11  
ABT16636/c  
ID ABT16636 standard; DNA; 282 BP.  
XX ABT16636;  
AC ABT16636;  
XX 03-APR-2003 (first entry)  
XX Artificial plant chromosome related oligo SEQ ID No 48.  
DE Plant artificial chromosome; PAC; transgenic plant; vaccine;  
XX blood factor; herbicide; stress; agronomical; nutrient quality;  
KW bacterial artificial chromosome; BAC; yeast artificial chromosome; YAC;  
XX ds.  
XX Unidentified.  
OS WC200296923-A1.  
XX 05-DEC-2002.  
PD 30-MAY-2002; 2002WO-US17451.  
XX 30-MAY-2001; 2001US-294687P.  
PR

```

PR 04-JUN-2001; 2001US-296329P.
XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.
PA (AGRI-) AGRISOMA INC.
XX
XX Perez C, Fabijanski SF, Perkins E;
XX WPI; 2003-140436/13.
XX
XX Producing artificial chromosome by introducing a nucleic acid into
XX plant cell, selecting artificial chromosome that has one or more repeat
XX regions with equivalent amounts of euchromatic and heterochromatic
XX nucleic acids -
XX
XX Disclosure; Page 263-264; 269pp; English.
XX
XX The invention relates to a novel method for producing plant artificial
XX chromosomes. The invention also relates to methods for targeting, methods
XX for insertion of heterologous DNA into plant artificial chromosomes, methods
XX for delivery of plant chromosomes to selected cells and tissues. The
XX isolated plant artificial chromosome (PAC) is useful for producing a
XX transgenic plant, which involves introducing the PAC into a plant cell.
XX The PAC comprises a heterologous nucleic acid encoding a gene product
XX such as enzymes, antisense RNA, tRNA, rDNA, structural proteins, marker
XX proteins, ligands, receptors, ribozymes, therapeutic proteins, and
XX biopharmaceutical proteins, vaccines, blood factors, antigens, hormones,
XX cytokines, growth factors, antibodies, or a product that provides for
XX resistance to diseases, insects, herbicides, or stress in a plant. The
XX heterologous nucleic acid optionally encodes a product that provides an
XX agronomically important trait in the plant, e.g. a product that alters
XX nutrient use and/or improves the nutrient quality of the plant. The
XX heterologous nucleic acid is contained within a bacterial artificial
XX chromosome (BAC) or a yeast artificial chromosome (YAC). This
XX polynucleotide sequence represents an oligo relating to the method for
XX producing plant artificial chromosomes of the invention.
XX
XX Sequence 282 BP; 79 A; 49 C; 39 G; 115 T; 0 other;
SQ
XX
XX Query Match 93.3%; Score 22.4; DB 25; Length 282;
XX Best Local Similarity 95.8%; Pred. No. 1.7;
XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 CAACTTAGTATCAAAAAGCTGAAC 24
XX |||||
XX Db 176 CAACTTAGTATCAAAAAGCTGAAC 153
XX
XX RESULT 12
XX AAF79770/c
XX ID AAF79770 standard; DNA; 610 BP.
XX
XX AC AAF79770;
XX
XX XX 29-MAY-2001 (first entry)
XX
XX Bacteriophage lambda attachment P region.
XX
XX DE Attachment P region; attP; recombination; marker gene removal; ds.
XX
XX KW Bacteriophage lambda.
XX
XX OS WO200121780-A2.
XX
XX PN 29-MAR-2001.
XX
XX PD 15-SEP-2000; 2000WO-GB03543.
XX
XX PF 17-SEP-1999; 99GB-0021937.
XX
XX PR (UYLE-) UNIV LEEDS.
XX
XX PA Meyer P, Zubko E;
XX
XX PI
XX
XX

```

---

```

DR WPI; 2001-266072/27.
XX
XX Removing a part of a transgene which has been integrated into a genome
XX comprises inducing intrachromosomal homologous recombination between
XX the attachment P regions of bacteriophage lambda flanking the transgene
XX -
XX
XX Claim 8; Fig 3D; 25pp; English.
XX
XX The present invention describes a method of removing a transgene marker
XX by flanking the transgene with a bacteriophage lambda attachment P (attP)
XX region and inducing homologous recombination between attP regions so that
XX the transgene is removed. This is useful in the production of transgenic
XX plants with less risk of inter-species transmission of marker genes,
XX which often encode proteins associated with, for example, herbicide and
XX antibiotic resistance. The present sequence is the attP coding region.
XX
XX Sequence 610 BP; 164 A; 108 C; 110 G; 228 T; 0 other;
SQ
XX
XX Query Match 93.3%; Score 22.4; DB 22; Length 610;
XX Best Local Similarity 95.8%; Pred. No. 1.8;
XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 CAACTTAGTATCAAAAAGCTGAAC 24
XX |||||
XX Db 300 CAACTTAGTATCAAAAAGCTGAAC 277
XX
XX RESULT 13
XX AAZ30709
XX ID AAZ30709 standard; cDNA; 1763 BP.
XX
XX AC AAZ30709;
XX
XX XX 05-JAN-2000 (first entry)
XX
XX DE Rat neuronal immediate early gene cDNA clone R280.
XX
XX KW Immediate early gene; IEG; neuron; brain; function; growth factor;
XX transcription factor; signal transduction; cytoskeletal protein;
XX metabolic enzyme; learning; memory; synaptic transmission; tolerance;
XX neuronal plasticity; ds.
XX
XX OS Rattus sp.
XX
XX PN WO9940225-A1.
XX
XX PD 12-AUG-1999.
XX
XX PF 05-FEB-1999; 99WO-US02462.
XX
XX PR 09-FEB-1998; 98US-0074135.
XX
XX PR 12-FEB-1998; 98US-0074518.
XX
XX PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
XX PA (BADI ) BASF-LYNX BIOSCIENCE AG.
XX
XX XX Worley PF, Lanhana A, Goetz B, Hiemisch H, Kuner R, Schesk S;
XX PI Nikolich K, Zhukovski E;
XX
XX WPI; 1999-590697/50.
XX
XX Novel genes and polypeptides, useful for treating conditions related to
XX a deficiency in NIEG responsiveness to a stimulus -
XX
XX Claim 1; Page 114-115; 134pp; English.
XX
XX This sequence represents rat neuronal immediate early gene (IEG) cDNA
XX clone R280. An IEG is a gene whose expression is rapidly
XX increased immediately following a stimulus e.g., neuronal stimulation.
XX Such neuronal IEGs have been found to encode a variety of proteins,
XX including transcription factors, cytoskeletal proteins, growth factors
XX and metabolic enzymes, as well as proteins involved in signal

```

transduction. The identification of neuronal IEGs and the proteins they encode may provide important information about the function of neurons in, for example, learning, memory, synaptic transmission, tolerance and neuronal plasticity. Neuronal IEGs, neuronal IEG protein products, cells expressing neuronal IEGs and antibodies against neuronal IEG proteins can be used to treat an animal with a deficiency in neuronal IEG responsiveness to stimuli, such that the effect of the deficiency is minimised. The deficiency may be a reduced or elevated level of expression of an IEG. The neuronal stimulus comprises a maximal electroconvulsive seizure and its effects influence learning or memory. The IEGs and protein products are useful in identifying compounds that modulate the expression or activity of IEG nucleic acids or proteins, respectively.

SQ Sequence 1763 BP; 478 A; 402 C; 452 G; 424 T; 7 other;

Query Match 93.3%; Score 22.4; DB 20; Length 1763;  
Best Local Similarity 95.8%; Pred. No. 1.9;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
|||||  
Db 270 CAACCTAGTATATAAAAAGCTGAAC 293

RESULT 14  
ACCA4716/c  
ID ACCA4716 standard; DNA; 4346 BP.

XX ACCA4716;

DT 29-MAY-2003 (first entry)

XX Plasmid pSV40193attPensePUR nucleotide sequence SEQ ID NO:113.

XX Chromosome-based platform; artificial chromosome; eukaryotic chromosome;  
KW att site; integrase; recombinase; ACes; gene therapy; transgenic animal;  
KW platform artificial chromosome expression system; gene; ds.

XX Bacteriophage lambda.

OS Synthetic.

FN WO200297059-A2.

XX 05-DEC-2002.

XX 30-MAY-2002; 2002WO-US17452.

XX 30-MAY-2001; 2001US-294758P.

XX 21-MAR-2002; 2002US-366891P.

XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.

PI Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E;  
PI Stewart S, Shellard J;

XX WPI; 2003-140461/13.

XX Novel eukaryotic chromosome comprising one or many att sites which  
PT permits site-directed integration in the presence of lambda-integrase,  
PT useful for site-specific recombination-directed integration of DNA of  
PT interest -

XX Example 3; Page 244-245; 272pp; English.

XX The present invention describes a eukaryotic chromosome (I) comprising  
CC one or several att sites, where an att site is heterologous to the  
CC chromosome, and permits site-directed integration in the presence of  
CC lambda-integrase. Also described: (1) a platform artificial chromosome  
CC expression system (ACes) (II) comprising several sites that participate  
CC in recombinase catalysed recombination; and (2) a method (M1) for  
CC introducing a heterologous nucleic acid into a platform artificial  
CC chromosome. (1) can be used in gene therapy. (M1) is useful for

CC introducing a heterologous nucleic acid molecule into a platform  
CC artificial chromosome, preferably an ACes. (II) is useful for producing a  
CC transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or  
CC mammal) by introducing (II) by cell fusion, lipid-mediated transfection,  
CC by a carrier system, microinjection, microcell fusion, electroporation,  
CC microprojectile bombardment or direct DNA transfer into an embryonic  
CC cell, preferably a stem cell or an embryo. (II) comprises a heterologous  
CC nucleic acid that encodes a therapeutic product which is useful for  
CC making a library of ACes comprising random portions of a genome. ACC44612  
CC to ACC44732 and ABP96650 to ABP96657 represent sequences used in the  
CC exemplification of the present invention.

SQ Sequence 4346 BP; 1022 A; 1208 C; 1094 G; 1022 T; 0 other;

Query Match 93.3%; Score 22.4; DB 25; Length 4346;  
Best Local Similarity 95.8%; Pred. No. 2;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTAGTATCAAAAAGCTGAAC 24  
|||||  
Db 4231 CAACCTAGTATATAAAAAGCTGAAC 4208

RESULT 15

ABT16615/c

ID ABT16615 standard; DNA; 4346 BP.

XX ABT16615;

DT 03-APR-2003 (first entry)

XX Artificial plant chromosome related plasmid DNA SEQ ID NO 26.

XX Plant artificial chromosome; PAC; transgenic plant; vaccine;

XX blood factor; herbicide; stress; agronomical; nutrient quality;

XX bacterial artificial chromosome; BAC; yeast artificial chromosome; YAC;

XX ds.

XX Unidentified.

XX WO200296923-A1.

XX 05-DEC-2002.

XX 30-MAY-2002; 2002WO-US17451.

XX 30-MAY-2001; 2001US-294687P.

XX 04-JUN-2001; 2001US-296329P.

XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.

PA (AGRI-) AGRISOMA INC.

XX Perez C, Fabijanski SF, Perkins E;

XX WPI; 2003-140436/13.

XX Producing artificial chromosome by introducing a nucleic acid into  
PT plant cell, selecting artificial chromosome that has one or more repeat  
PT regions with equivalent amounts of euchromatic and heterochromatic  
PT nucleic acids -

XX Example 19; Page 255-256; 269pp; English.

XX The invention relates to a novel method for producing plant artificial  
CC chromosomes. The invention also relates to methods for targeting  
CC insertion of heterologous DNA into plant artificial chromosomes, methods  
CC for delivery of plant chromosomes to selected cells and tissues. The  
CC isolated plant artificial chromosome (PAC) is useful for producing a  
CC transgenic plant, which involves introducing the PAC into a plant cell.  
CC The PAC comprises a heterologous nucleic acid encoding a gene product  
CC such as enzymes, antisense RNA, tRNA, rDNA, structural proteins, marker  
CC proteins, ligands, receptors, ribozymes, therapeutic proteins, and  
CC biopharmaceutical proteins, vaccines, blood factors, antigens, hormones,

CC cytokines, growth factors, antibodies, or a product that provides for  
CC resistance to diseases, insects, herbicides, or stress in a plant. The  
CC heterologous nucleic acid optionally encodes a product that provides an  
CC agronomically important trait in the plant, e.g. a product that alters  
CC nutrient use and/or improves the nutrient quality of the plant. The  
CC heterologous nucleic acid is contained within a bacterial artificial  
CC chromosome (BAC) or a yeast artificial chromosome (YAC). This  
CC polynucleotide sequence represents the DNA of a plasmid used in the  
CC method of the invention.  
XX  
SQ Sequence 4346 BP; 1022 A; 1208 C; 1094 G; 1022 T; 0 other;  
Query Match 93.3%; Score 22.4; DB 25; Length 4346;  
Best Local Similarity 95.8%; Pred. No. 2;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CAACTTAGTATCAAAAAGCTGAAC 24  
|||||  
Db 4231 CAACTTAGTATCAAAAAGCTGAAC 4208  
Search completed: December 3, 2003, 12:23:36  
Job time : 115.843 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2003, 13:18:51 ; Search time 107.107 Seconds  
(without alignments)  
744.732 Million cell updates/sec

Title: US-10-082-772-3

Perfect score: 24

Sequence: 1 gttcagcttttgataactaagtgtg 24

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	100.0	24	14	US-10-082-772-3
2	24	100.0	24	14	US-10-082-772-4
3	22.4	93.3	27	9	US-09-732-914-2
4	22.4	93.3	27	14	US-10-151-690-26
5	22.4	93.3	282	14	US-10-161-403-72
6	22.4	93.3	1763	12	US-09-244-805-57
7	22.4	93.3	1763	12	US-09-245-277-57
8	22.4	93.3	4346	14	US-10-161-403-113
9	21.2	88.3	25	9	US-09-855-797A-43
10	21.2	88.3	25	10	US-09-907-900-43
11	21.2	88.3	25	10	US-09-907-719-43
12	21.2	88.3	25	12	US-09-985-448-43
13	21.2	88.3	25	12	US-10-300-892-43
14	20.8	86.7	27	9	US-09-732-914-4
15	20.8	86.7	27	14	US-10-151-690-28
16	20.4	85.0	25	9	US-09-855-797A-42

17	20.4	85.0	25	10	US-09-907-900-42	Sequence 42, Appl
18	20.4	85.0	25	10	US-09-907-719-42	Sequence 42, Appl
19	20.4	85.0	25	12	US-09-985-448-42	Sequence 42, Appl
20	20.4	85.0	25	12	US-10-300-892-42	Sequence 42, Appl
21	19.4	80.8	21	11	US-09-981-803-4	Sequence 4, Appl
22	19.4	80.8	21	14	US-10-252-384-7	Sequence 7, Appl
23	19.4	80.8	21	14	US-10-161-403-33	Sequence 33, Appl
24	19.4	80.8	48	9	US-09-732-914-70	Sequence 70, Appl
25	19.2	80.0	25	9	US-09-732-914-32	Sequence 32, Appl
26	19.2	80.0	25	9	US-09-732-914-36	Sequence 36, Appl
27	19.2	80.0	25	9	US-09-855-797A-15	Sequence 15, Appl
28	19.2	80.0	25	10	US-09-907-900-15	Sequence 15, Appl
29	19.2	80.0	25	10	US-09-907-719-15	Sequence 15, Appl
30	19.2	80.0	25	11	US-09-432-085-15	Sequence 15, Appl
31	19.2	80.0	25	12	US-09-985-448-15	Sequence 15, Appl
32	19.2	80.0	25	12	US-10-300-892-15	Sequence 10, Appl
33	19.2	80.0	25	14	US-10-055-001A-10	Sequence 15, Appl
34	19.2	80.0	25	14	US-10-058-293-15	Sequence 15, Appl
35	19.2	80.0	25	14	US-10-162-879-15	Sequence 15, Appl
36	19.2	80.0	25	14	US-10-161-403-55	Sequence 55, Appl
37	19.2	80.0	25	14	US-10-151-690-56	Sequence 56, Appl
38	19.2	80.0	25	14	US-10-151-690-60	Sequence 60, Appl
39	19.2	80.0	27	9	US-09-732-914-6	Sequence 6, Appl
40	19.2	80.0	27	9	US-09-732-914-30	Sequence 30, Appl
41	19.2	80.0	27	9	US-09-732-914-34	Sequence 34, Appl
42	19.2	80.0	27	14	US-10-151-690-30	Sequence 30, Appl
43	19.2	80.0	27	14	US-10-151-690-54	Sequence 54, Appl
44	19.2	80.0	27	14	US-10-151-690-58	Sequence 58, Appl
45	19.2	80.0	4470	14	US-10-151-690-21	Sequence 21, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-082-772-3  
; Sequence 3, Application US/10082772  
; Publication No. US20030027337A1  
; GENERAL INFORMATION:

; APPLICANT: DROGE, PETER  
; APPLICANT: CHRIST, NICOLE  
; APPLICANT: LOEBACH, ELKE  
; TITLE OF INVENTION: SEQUENCE-SPECIFIC DNA RECOMBINATION IN EUKARYOTIC CELLS  
; FILE REFERENCE: DEBE:008US  
; CURRENT APPLICATION NUMBER: US/10/082,772  
; CURRENT FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: PCT/DE 00/02947  
; PRIOR FILING DATE: 2000-08-29  
; PRIOR APPLICATION NUMBER: DE 199 41 186.7  
; PRIOR FILING DATE: 1999-08-30  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Primer  
US-10-082-772-3

Query Match 100.0%; Score 24; DB 14; Length 24;

Best Local Similarity 100.0%; Pred. No. 0.31; 0; Indels 0; Gaps 0;  
Matches 24; Conservative 0; Mismatches 0;

OY 1 GTTCAGCTTTTGATACTAAGTTG 24

DB 1 GTTCAGCTTTTGATACTAAGTTG 24

#### RESULT 2

US-10-082-772-4/c

; Sequence 4, Application US/10082772

```

; Publication No. US20030027337A1
; GENERAL INFORMATION:
; APPLICANT: DROGE, PETER
; APPLICANT: CHRIST, NICOLE
; APPLICANT: LORRACH, ELKE
; TITLE OF INVENTION: SEQUENCE-SPECIFIC DNA RECOMBINATION IN EUKARYOTIC CELLS
; FILE REFERENCE: DEBE:0080US
; CURRENT APPLICATION NUMBER: US/10/082.772
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: PCT/DE 00/02947
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: DE 199 41 186.7
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
US-10-082-772-4

Query Match 100.0%; Score 24; DB 14; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTACTAAGTTG 24
Db 24 GTTCAGCTTTTGTACTAAGTTG 1

RESULT 3
US-03-732-914-2
; Sequence 2, Application US/09732914
; Patent No. US20020007051A1
; GENERAL INFORMATION:
; APPLICANT: Cheo, David
; APPLICANT: Brasch, Michael A.
; APPLICANT: Temple, Gary F.
; APPLICANT: Hartley, James L.
; APPLICANT: Byrd, Devon R.N.
; TITLE OF INVENTION: Use of Multiple Recombination Sites with Unique Specificity in
; TITLE OF INVENTION: Recombinational Cloning
; FILE REFERENCE: 0942.501002
; CURRENT APPLICATION NUMBER: US/09/732.914
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/169,983
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: US 60/188,020
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 27
; TYPE: DNA
; ORGANISM: attp0
US-09-732-914-2

Query Match 93.3%; Score 22.4; DB 9; Length 27;
Best Local Similarity 95.8%; Pred. No. 1.6;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTACTAAGTTG 24
Db 1 GTTCAGCTTTTGTACTAAGTTG 24

RESULT 4
US-10-151-690-26
; Sequence 26, Application US/10151690
; Publication No. US20030124555A1

```

```

; GENERAL INFORMATION:
; APPLICANT: BRASCH, MICHAEL A.
; APPLICANT: CHEO, DAVID
; APPLICANT: LI, XIAO
; APPLICANT: ESPOSITO, DOMINIC
; APPLICANT: BYRD, DEVON R.N.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR USE IN ISOLATION OF NUCLEIC ACID MOLE
; FILE REFERENCE: 0942.5120001
; CURRENT APPLICATION NUMBER: US/10/151.690
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 10/151.690
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/291.973
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 27
; TYPE: DNA
; ORGANISM: attp0
US-10-151-690-26

Query Match 93.3%; Score 22.4; DB 14; Length 27;
Best Local Similarity 95.8%; Pred. No. 1.6;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTACTAAGTTG 24
Db 1 GTTCAGCTTTTGTACTAAGTTG 24

RESULT 5
US-10-161-403-72
; Sequence 72, Application US/10161403
; Publication No. US20030119104A1
; GENERAL INFORMATION:
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
; FILE REFERENCE: 24601-420
; CURRENT APPLICATION NUMBER: US/10/161.403
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: attp
US-10-161-403-72

Query Match 93.3%; Score 22.4; DB 14; Length 282;
Best Local Similarity 95.8%; Pred. No. 2.6;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTACTAAGTTG 24
Db 153 GTTCAGCTTTTGTACTAAGTTG 176

RESULT 6
US-09-244-805-57/c

```

Sequence 57, Application US/09244805  
Publication No. US20030203840A1  
GENERAL INFORMATION:  
APPLICANT: Worley, Paul F.  
APPLICANT: Lananan, Anthony  
APPLICANT: Goetz, Bernard  
APPLICANT: Heimsch, Holger  
APPLICANT: Kuner, Rohini  
APPLICANT: Scheek, Sigrid  
APPLICANT: Nikolich, Karoly  
APPLICANT: Zhukovski, Eugene  
TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE  
TITLE OF INVENTION: THEREFOR  
FILE REFERENCE: 10496/004001  
CURRENT APPLICATION NUMBER: US/09/244,805  
CURRENT FILING DATE: 1999-02-05  
EARLIER APPLICATION NUMBER: 60/074,518  
EARLIER FILING DATE: 1998-02-12  
EARLIER APPLICATION NUMBER: 60/074,135  
EARLIER FILING DATE: 1998-02-06  
NUMBER OF SEQ ID NOS: 62  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 57  
LENGTH: 1763  
TYPE: DNA  
ORGANISM: Eukaryote  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(1763)  
OTHER INFORMATION: Y = C or T  
OTHER INFORMATION: n = A,T,C or G  
US-09-244-805-57

Query Match 93.3%; Score 22.4; DB 12; Length 1763;  
Best Local Similarity 95.8%; Pred. No. 3.7;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGCATCTAAGTTG 24  
Db 293 GTTCAGCTTTTATATACTAAGTTG 270

RESULT 7

US-09-245-277-57/c  
Sequence 57, Application US/09245277  
Publication No. US20030211984A1  
GENERAL INFORMATION:  
APPLICANT: Worley, Paul F.  
APPLICANT: Lananan, Anthony  
TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE  
TITLE OF INVENTION: THEREFOR  
FILE REFERENCE: JHU1530-3  
CURRENT APPLICATION NUMBER: US/09/245,277  
CURRENT FILING DATE: 1999-02-05  
PRIOR APPLICATION NUMBER: 60/074,518  
PRIOR FILING DATE: 1998-02-12  
PRIOR APPLICATION NUMBER: 60/074,135  
PRIOR FILING DATE: 1998-02-06  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 57  
LENGTH: 1763  
TYPE: DNA  
ORGANISM: Eukaryote  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(1763)  
OTHER INFORMATION: Y = C or T  
OTHER INFORMATION: n = A,T,C or G  
US-09-245-277-57

Query Match 93.3%; Score 22.4; DB 12; Length 1763;  
Best Local Similarity 95.8%; Pred. No. 3.7;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GTTCAGCTTTTGCATCTAAGTTG 24  
Db 293 GTTCAGCTTTTATATACTAAGTTG 270

RESULT 8

US-10-161-403-113  
Sequence 113, Application US/10161403  
Publication No. US20030119104A1  
GENERAL INFORMATION:  
APPLICANT: Perkins, Edward  
APPLICANT: Lindenbaum, Michael  
APPLICANT: Greene, Amy  
APPLICANT: Leung, Josephine  
APPLICANT: Fleming, Elena  
APPLICANT: Stewart, Sandra  
APPLICANT: Shellard, Joan  
TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
TITLE OF INVENTION: THEREFOR  
FILE REFERENCE: 24601-420  
CURRENT APPLICATION NUMBER: US/10/161,403  
CURRENT FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: 60/294,758  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 60/366,891  
PRIOR FILING DATE: 2002-03-21  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 113  
LENGTH: 4346  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: pSV40-193AttpeensePur Plasmid  
US-10-161-403-113

Query Match 93.3%; Score 22.4; DB 14; Length 4346;  
Best Local Similarity 95.8%; Pred. No. 4.4;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGCATCTAAGTTG 24  
Db 4208 GTTCAGCTTTTATATACTAAGTTG 4231

RESULT 9

US-09-855-797A-43  
Sequence 43, Application US/09855797A  
Patent No. US20020094574A1  
GENERAL INFORMATION:  
APPLICANT: Hartley, James L.  
APPLICANT: Brasch, Michael A.  
APPLICANT: Temple, Gary F.  
APPLICANT: Fox, Donna K.  
TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having  
TITLE OF INVENTION: Recombination Sites  
FILE REFERENCE: 0942.2850008  
CURRENT APPLICATION NUMBER: US/09/855,797A  
CURRENT FILING DATE: 2001-05-16  
PRIOR APPLICATION NUMBER: 09/296,281  
PRIOR FILING DATE: 1999-04-22  
PRIOR APPLICATION NUMBER: US 60/065,930  
PRIOR FILING DATE: 1997-10-24  
NUMBER OF SEQ ID NOS: 60  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 43  
LENGTH: 25  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: recombination  
US-09-855-797A-43

; OTHER INFORMATION: products  
US-09-855-797A-43

```
Query Match      88.3%; Score 21.2; DB 9; Length 25;
Best Local Similarity 83.3%; Pred. NO. 5.6;
Matches 20; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 GTTCAGCTTTTGTGATACTAAGTTG 24  
|||||:|:|:|:|  
Db 1 GTTCAGCTTTTGTGATACTAAGTTG 24

## RESULT 10

```

US-09-907-900-43
; Sequence 43, Application US/09907900
; Patent No. US20020172997A1
; GENERAL INFORMATION:
; APPLICANT: Hartley, James L.
; APPLICANT: Brasch, Michael A.
; APPLICANT: Temple, Gary F.
; APPLICANT: Fox, Donna K.
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having
; TITLE OF INVENTION: Recombination Sites
; FILE REFERENCE: 0942.2850004
; CURRENT APPLICATION NUMBER: US/09/907,900
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 09/177,387
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: recombination
; OTHER INFORMATION: products
US-09-907-900-43

```

Query Match 88.3%; Score 21.2; DB 10; Length 25;  
Best Local Similarity 83.3%; Pred. No. 5.6;  
Matches 20: Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCAGCTTTTGTGATACTAAGTTG 24  
||| ||| : ||| : |||  
Db 1 GTTCAGCTTTTGTGATACTAAGTTG 24

## RESULT 11

```

US-09-907-719-43
; Sequence 43, Application US/09907719
; Publication No. US20020192819A1
; GENERAL INFORMATION:
; APPLICANT: Hartley, James L.
; APPLICANT: Brasch, Michael A.
; APPLICANT: Temple, Gary F.
; APPLICANT: Fox, Donna K.
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having
; TITLE OF INVENTION: Recombination Sites
; FILE REFERENCE: 0942.2850004
; CURRENT APPLICATION NUMBER: US/09/907,719
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US/09/177,387
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: recombination
; OTHER INFORMATION: products

```

US-09-907-719-43

```
Query Match      88.3%; Score 21.2; DB 10; Length 25;
Best Local Similarity 83.3%; Pred. No. 5.6;
Matches 20; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 GTTCAGCTTTTGTACTAAGTTG 24  
|||||:|:|:|  
Dp 1 GTTCAGCTTTTTRTACWAAAGTTG 24

## RESULT 12

```

US-09-985-448-43
; Sequence 43, Application US/09985448
; Publication No. US20030157716A1
; GENERAL INFORMATION:
; APPLICANT: Hartley, James L.
; APPLICANT: Brasch, Michael A.
; APPLICANT: Temple, Gary F.
; APPLICANT: Fox, Donna K.
; TITLE OF INVENTION: Recombinational Cloning Using Nucleic Acids Having
; TITLE OF INVENTION: Recombination Sites
; FILE REFERENCE: 0942.2850004
; CURRENT APPLICATION NUMBER: US/09/985,448
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US/09/177,387
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: US 60/065,930
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 43
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown: recombination
; OTHER INFORMATION: products
; US-09-985-448-43

```

Query Match	88.3%	Score 21.2;	DB 12;	Length 25;
Best Local Similarity	83.3%	Pred. NO. 5.6;		
Matches 20: Conservative	3;	Mismatches 1;	Indels 0;	Gaps 0;

QY 1 GTTCAGCTTTTGTGATACTAAGTTG 24  
||||| : |||  
db 1 GTTCAGCTTTTGTGATACTAAGTTG 24  
||||| : |||

## RESULT 13

```

US-10-300-892-43
; Sequence 43, Application US/10300892
; Publication No. US20030175970A1
; GENERAL INFORMATION:
; APPLICANT: Hartley, James L.
; APPLICANT: Brasch, Michael A.
; APPLICANT: Temple, Gary F.
; APPLICANT: Fox, Donna K.
; TITLE OF INVENTION: Recombination Cloning Using Nucleic Acids Having
; TITLE OF INVENTION: Recombination Sites
; FILE REFERENCE: 0942.2850004
; CURRENT APPLICATION NUMBER: US/10/300,892
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/907,719
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US/09/177,387
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Unknown

```



FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: recombination  
OTHER INFORMATION: products  
US-10-300-892-43

Query Match 88.3%; Score 21.2; DB 12; Length 25;  
Best Local Similarity 83.3%; Pred. No. 5.6;  
Matches 20; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGATCTAAGTTG 24  
|||||:|||||  
Db 1 GTTCAGCTTTTTRTACWAAGTTG 24

## RESULT 14

US-09-732-914-4  
Sequence 4, Application US/09732914  
Patent No. US20020007051A1  
GENERAL INFORMATION:  
APPLICANT: Cheo, David  
APPLICANT: Brasch, Michael A.  
APPLICANT: Temple, Gary F.  
APPLICANT: Hartley, James L.  
APPLICANT: Byrd, Devon R.N.  
TITLE OF INVENTION: Use of Multiple Recombination Sites with Unique Specificity in  
FILE REFERENCE: 0942.5010002  
CURRENT APPLICATION NUMBER: US/09/732,914  
CURRENT FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 60/169,983  
PRIOR FILING DATE: 1999-12-10  
PRIOR APPLICATION NUMBER: US 60/188,020  
PRIOR FILING DATE: 2000-03-09  
NUMBER OF SEQ ID NOS: 140  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 27  
TYPE: DNA  
ORGANISM: attro  
US-09-732-914-4

Query Match 86.7%; Score 20.8; DB 9; Length 27;  
Best Local Similarity 91.7%; Pred. No. 8.5;  
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCAGCTTTTGATCTAAGTTG 24  
|||||:|||||  
Db 1 GTTCAGCTTTTATCTACTTG 24

## RESULT 15

US-10-151-690-28  
Sequence 28, Application US/10151690  
Publication No. US20030124555A1  
GENERAL INFORMATION:  
APPLICANT: BRASCH, MICHAEL A.  
APPLICANT: CHEO, DAVID  
APPLICANT: LI, XIRO  
APPLICANT: ESPOSITO, DOMINIC  
APPLICANT: BYRD, DEVON R.N.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR USE IN ISOLATION OF NUCLEIC ACID MOL  
FILE REFERENCE: 0942.5120001  
CURRENT APPLICATION NUMBER: US/10/151,690  
CURRENT FILING DATE: 2002-05-21  
PRIOR APPLICATION NUMBER: US 10/151,690  
PRIOR FILING DATE: 2002-05-21  
PRIOR APPLICATION NUMBER: US 60/291,973  
PRIOR FILING DATE: 2001-05-21  
NUMBER OF SEQ ID NOS: 64  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 28  
LENGTH: 27  
TYPE: DNA

ORGANISM: attro  
US-10-151-690-28

Query Match 86.7%; Score 20.8; DB 14; Length 27;  
Best Local Similarity 91.7%; Pred. No. 8.5;  
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GTTCAGCTTTTGATCTAAGTTG 24  
|||||:|||||  
Db 1 GTTCAGCTTTTATCTACTTG 24

Search completed: December 3, 2003, 17:22:25  
Job time : 108.107 secs